ISTANBUL TECHNICAL UNIVERSITY ★ INFORMATICS INSTITUTE

UNSUPERVISED ACTIVE LEARNING FOR VIDEO ANNOTATION

M.Sc. THESIS

Emre DEMİR

Informatics Institute

Computer Sciences Department

MAY 2015

ISTANBUL TECHNICAL UNIVERSITY ★ INFORMATICS INSTITUTE

UNSUPERVISED ACTIVE LEARNING FOR VIDEO ANNOTATION

M.Sc. THESIS

Emre DEMİR
(704101003)

Informatics Institute

Computer Sciences Department

Thesis Advisor: Prof. Zehra ÇATALTEPE

MAY 2015

İSTANBUL TEKNİK ÜNİVERSİTESİ ★ BİLİŞİM ENSTİTÜSÜ

VİDEO ETİKETLEME İÇİN AKTİF ÖĞRENME

YÜKSEK LİSANS TEZİ

Emre DEMİR
(704101003)

Bilişim Enstitütüsü

Bilgisayar Bilimleri Bölümü

Tez Danışmanı: Prof. Zehra ÇATALTEPE

MAYIS 2015

Emre DEMİR, a M.Sc. student of ITU Informatics Institute 704101003 successfully defended the thesis entitled **"UNSUPERVISED ACTIVE LEARNING FOR VIDEO ANNOTATION"**, which he/she prepared after fulfilling the requirements specified in the associated legislations, before the jury whose signatures are below.

Thesis Advisor :	**Prof. Zehra ÇATALTEPE** Istanbul Technical University	.....
Jury Members :	**Assist. Prof. Yusuf YASLAN** Istanbul Technical University	.....
	Assist. Prof. Arzucan ÖZGÜR Boğaziçi University	.....

Date of Submission : **May 2015**
Date of Defense : **28 May 2015**

To my wife Melis,

FOREWORD

I would like to thank Prof.Dr. Zehra Çataltepe for her assistance, motivation and involvement in every step throughout the thesis.

I would like to thank Umit Ekmekci and Mateusz Budnik for their help, ideas and collaboration.

I thank to my beloved wife Melis Özgür Çetinkaya Demir for her encouragement and support.

Work in this paper is partially supported by the Scientific and Technological Research Council of Turkey (TUBITAK) project 112E176.

May 2015 Emre DEMİR

TABLE OF CONTENTS

Page

FOREWORD ..... **ix**
TABLE OF CONTENTS ..... **xi**
ABBREVIATIONS ..... **xiii**
LIST OF TABLES ..... **xv**
LIST OF FIGURES ..... **xvii**
SUMMARY ..... **xix**
ÖZET ..... **xxi**
1. INTRODUCTION **1**
1.1 Thesis Organization 2
2. NOTATION AND DATASET **3**
2.1 Notation 3
2.2 Dataset 3
3. RELATED WORK **7**
3.1 Active Learning 7
3.1.1 Active Learning Scenarios 8
3.1.1.1 Query Synthesis 9
3.1.1.2 Stream Based Selective Sampling 9
3.1.1.3 Pool-Based Sampling 10
3.1.2 Sampling Strategies 10
3.1.2.1 Uncertainty Sampling 10
Least Confident 11
Margin 11
Entropy 11
3.1.2.2 Query by Disagreement 12
3.1.2.3 Query by Committee 12
3.2 Cluster Matching Metrics 13
4. METHODOLOGY **19**
4.1 Clustering Algorithms 19
4.1.1 Agglomerative Clustering 21
4.1.2 K-Medoid Clustering 21
4.2 Query Instance Selection Strategies 23
4.2.1 Cluster Selection 25
4.2.1.1 Big Cluster First 25
4.2.1.2 The Most Disagree Selection 25
4.2.1.3 Hybrid Cluster Selection 26

4.2.2 Instance Selecting Strategy 27
5. EXPERIMENTS AND RESULTS 29
5.1 Evaluation Metrics 29
5.2 Results 30
6. CONCLUSIONS AND RECOMMENDATIONS 37
REFERENCES 39
CURRICULUM VITAE 44

ABBREVIATIONS

BCPS : Big Cluster Pair First Selection
BCS : Big Cluster First Selection
HHC : Hard Hybrid Certainty Selection
HHU : Hard Hybrid Uncertainty Selection
MDS : Most Disagree Selection
DS : Disagreement Score
DSL : Disagreement Score List
HYB : Soft Hybrid Selection
H-HYB : Hard Hybrid Selection
KL : Kullback-Leiber Divergence
OCR : Optical Character Recognition
QDB : Query By Disagreement
QDC : Query By Committee
SHS : Soft Hybrid Score
VI : Variation of Information
ST : Speaker Track
FT : Face Track
ON : Overlaid Name
SS : Speaker Score with Speaker Annotation
FS : Face Score with Speaker Annotation
FF : Face Score with Face Annotation
SF : Speaker Score with Face Annotation

LIST OF TABLES

Page

Table 2.1 : The video list and the size of distance matrices which are inputs for our algorithms. 6

Table 4.1 : A sample cluster pair similarity matrix for agglomerative and K-Medoid algorithms. 22

LIST OF FIGURES

Page

Figure 2.1 : Example frames for face segmentation and OCR [1]. 4
Figure 3.1 : Detecting safe shaped food threshold θ with binary search. 8
Figure 4.1 : The architecture of the system. 20
Figure 4.2 : The illustration of Agglomerative Clustering. 21
Figure 4.3 : Pseudo-code of Agglomerative Clustering 21
Figure 4.4 : Pseudo-code of K-Medoid algorithm. 22
Figure 4.5 : The stable matching algorithm. 24
Figure 5.1 : Face score with face annotation. 31
Figure 5.2 : Speaker score with face annotation. 31
Figure 5.3 : Face score with with speaker annotation. 31
Figure 5.4 : Speaker score with with speaker annotation. 31
Figure 5.5 : Face score with face annotation 33
Figure 5.6 : Speaker score with face annotation. 33
Figure 5.7 : Face score with with speaker annotation. 33
Figure 5.8 : Speaker score with with speaker annotation. 33
Figure 5.9 : Face score with face annotation (lower is better). 34
Figure 5.10 : Speaker score with face annotation (lower is better). 34
Figure 5.11 : Face score with with speaker annotation (lower is better) 34
Figure 5.12 : Speaker score with with speaker annotation (lower is better). 34

UNSUPERVISED ACTIVE LEARNING FOR VIDEO ANNOTATION

SUMMARY

When annotating complex multimedia data like videos, a human expert usually annotates them manually. Even tough manual annotation achieves accurate results, it is a labor-intensive and time-consuming process. On the other hand, computational methods can annotate mass video data for indexing and searching with any or almost no help from human experts effortlessly and faster but they are probably more error prone solutions. The tradeoff between the costs in terms of labor, time and accuracy reveals Active Learning as a natural outcome. Active learning is one of the semi-supervised machine learning methods that benefits from the strongest properties of both manual and computational methods. In an active learning cycle, a learner algorithm discovers the underlying patterns in data and queries the human experts interactively for some informative decision points. It is used when labeled instances are insufficient and acquiring new labels is expensive or especially when unlabeled instances are abundant. In this study, we introduce an unsupervised active learning cycle structure in a flow, which includes clustering, stable matching between the created clusters, various unsupervised selection strategies for selecting the most uncertain and the most certain instances and querying the human annotators. We propose two new cluster selection methods, namely Most Disagreement Selection (MDS) and Hybrid Set Selection (HS), which is a hybrid of MDS and Big Cluster First [2] methods. For MDS and HS, we adopt the "Stable Marriage Problem" solution, in which a stable marriage problem is transformed into a cluster matching problem. We work on REPERE [1] video dataset, which is created for the problem of person identification in videos. Our study aims to identify who is speaking and who is on screen by using multi-modal data. We have evaluated the performance of selection strategies over active learning cycles using multimodality on 28 videos from 7 different TV programs. Each video has three different similarity matrices namely face-to-face, speech-to-speech and face-to-speech. We have run four experiments with regard to matrices in this order: face score for face track annotation, face score for speaker track annotation, speaker score for speaker track annotation and speaker score for speaker annotation.

VİDEO ETİKETLEME İÇİN AKTİF ÖĞRENME

ÖZET

Günümüzde dijital video ve fotoğraf içeriklerinin kolay üretilebilmesi ve bu üretilen videoların internet üstünden medya, sosyal medya ve video paylaşım siteleri gibi kanallar üstünden internet sitelerine son kullanıcıların da çok kolay yükleme yapabilmesi nedeniyle her geçen gün internet ortamında muazzam derecede içerik oluşmasına neden olmaktadır. Video verilerinin sürekli artması ve bu verilere kullanıcıların çok kolay erişmesi, bilgiye ulaşmayı oldukça arttırmaktadır ancak video verisinin artması ve videoların içeriklerinin etiketlenmesinin zor olmasından dolayı aynı zamanda doğru videolara erişim, istenilen videoyu arama, istenilen videoların arşivlenmesi vb. işlemleri zor kılmaktadır. Bunun nedenlerinden başında video verisini üreten ve/veya video verisini kayıt eden kullanıcıların video içeriklerini etiketlememesi veya eksik etiketlemesidir. Bu durum etiketlenmemiş videoların aranıp bulunmasını oldukça zorlaştırmaktadır. Ayrıca etiketlenmemiş, eksik etiketlenlemiş veya yanlış etiketlenmiş video verileri aynı zamanda arama veya arşivleme gibi işlemlerde hatalı ve gürültülü sonuçlara neden olmaktadır.

Kullanıcıların videoları etiketlememesi veya düzgün etiketlememesinin nedenlerinden birisi videoların içeriklerinin etiketlenmesi işleminin zor olmasıdır. Tüm videolar baştan sona izlenip içeriklerinin ayrıntılı olarak etiketlenmesi gerekmektedir. Bu etiketler video içeriğinde hangi insanların olduğu gibi etiketler olabilirken, videoda kayıt edilen genel bir olay(basketbol, yangın videosu) veya spesifik bir olay da olabilir, dünya kupası finali gibi. Kullanıcıların etiketlemedikleri veya eksik etiketleme yaptıları durumlarda video verisini etiketlenebilmesi için otomatik video etiketleme konusunda araştırma ve geliştirme yapılmaktadır.

Bu çalışmada otomatik video etiketleme yapacak makine öğrenmesi algoritmalarının eğitimi sırasında gereksinim duyduğu etiketlenmiş video verisinin azaltılmasına odaklanılmıştır. Denetimli makine öğrenmesi temelli otomatik video etiketleme sistemleri, öğrenme safası öncesinde insan uzmanlar tarafından etiketlenmiş belli sayıda örnek kullanır. Denetimli makine öğrenme algoritmalarının eğitim sürecince kullanacağı örnekler genellikle rastgele seçilir ve insan uzman tarafından etiketleme işlemi yapılır. Otomatik etiketleme sistemlerinin öğrenme safası için gereksinim duydukları bu etiketlenmiş örneklerin bir insan tarafından etiketlenmesi süreci çok uzun sürmektedir. çoğu zaman etiketlenecek video kare kare incelenip her kare etiketlenir. Etiketleme işleminin uzun sürmesi nedeniyle makine örenmesi algoritmalarının öğrenim sürecinde daha az etiketlenmiş örneğe gereksinim duyan yöntemler araştırılmakta ve geliştirilmektedir. Bu yöntemlerden bir tanesi yarı denetimli bir yöntem olan aktif öğrenmedir.

Denetimli öğrenme yöntemlerinde, hipotezin gereksinim duyduğu kadar örnek bir uzman tarafından etiketlenir ve bu örnekler tüm örnek havuzundan rastgele seçilir. Buna karşılık yarı denetimli yöntem olan aktif öğrenme algoritmaları çok az sayıda rastgele etiketlenmiş veri ile öğrenim sürecine başlar ve denetimli öğrenmeye göre daha az sayıda etiketlenmiş örnek ile denetimli makine öğrenmesi yöntemiyle aynı başarı seviyesine gelmeyi amaçlar. Bunun başarılabilmesi örnekleri rasgele seçmek yerine, sınıflandırıcı için en çok bilgi taşıyan örneklerin bulunmasına bağlıdır. En çok bilgi taşıyan örnekler teorik olarak karar sınırlarının çevresinde bulunur. İkili bir sınıflandırma problemi için anlatacak olursak, karar sınırlarının üstünde yer alan bir örneğin her iki sınıfa da ait olma olasılığı 0.5 dir. Olasılık 0.5 olduğu durumda entropi en yüksek değerini alır ve bu noktaların bulunması etiketlenecek örnek sayısında düşüşe neden olması beklenir.Bu noktada aktif öğrenme çözülmeye çalışılan problem için en çok bilgi taşıyan örnekleri bulmaya çalışır, teoride karar sınırlarının etrafında yer alan örnekler en çok bilgiye sahip olması beklensede önceki çalışmalar göstermiştir ki bazı hipotezler de karar sınırının en uzağında yer alan örnekler daha çok bilgi taşıyabilmektedir.

Biz bu çalışmada video etiketleme için aktif öğrenme yöntemini denetimsiz öğrenme problemi üstünde uyguladık. Böylece otomatik video etiketleme sisteminin rastgele etiketleme ile aynı seviyede başarıyı daha az sayıda etiketlenmiş video ile elde edilmesi amaçlandı.

Aktif öğrenme fikiri 1980′li yıllarda çıkmış olup, uygulamarı günümüze kadar yaygınlaşmamıştır. Günümüzde aktif öğrenme ile ilgili çalışmalar hızlanmıştır ve bu hızlanmanın ana nedeni büyük veriler üstünde öğrenme algoritmalarının öğrenme safasını hızlandırmaktır. Bu çalışmada önerdiğimiz yöntem denetimsiz bir makine öğrenmesi olan kümeleme üstünde çalışmaktadır. çalışmanın detaylarında anlatıldığı gibi aktif öğrenme algoritmaları genellikle denetimli yöntemler için geliştirilmiş ve denetimli yöntemler üstünde uygulamaları daha yaygındır. Bu çalışma sırasında denetimsiz öğrenme için aktif öğrenme tekniği geliştirilik problemleri çözülmüştür.

Önerilen yöntem denetimli aktif öğrenme yöntemi olan ”anlaşmazlığa göre sorgulama” (query by disagreement) ve ”belirsizi sorgulama” (uncertainty sampling) yöntemlerinin denetimsiz öğrenme için geliştirilmişidir.

Anlaşmazlığa göre sorgulama denetimli yöntemlerde uygulaması şu şekilde tanımlanır; aynı veri kümesi kullanılarak iki farklı sınıflandırıcı eğitilir, bu iki farklı sınıflandırıcı aynı tip sınıflandırıcının fazla eğtilmiş bir hipotezi ve az eğitilmiş başka bir hipotezi olabileceği gibi iki farklı tip sınıflandırıcı da olabilir. İki farklı sınıflandırıcı eğiltildikten sonra bu iki sınıflandırıcı etiketlenmemiş örnekleri ayrı ayrı etiketler ve yani her bir örneğin farklı sınıflandırıcılardan gelen iki adet etiketi olur. Bir örneğe atanan iki farklı etiket birbirinden farklı ise yani sınıflandırıcılar anlaşmazlığa düşmüş durumdaysa, bu örnek etiketlenmek üzere insan uzamana sorulur. Böylece örneklerin dağlımında sınıflar arasındaki karar sınırına en yakın örneklerin bulunması amaçlanır çünkü iki sınıflandırıcının en çok anlaşmazlığa düştüğü noktalar karar sınırına yakın olanların olması beklenir.

Bir başka aktif öğrenme yöntemi en basit ve en çok kullanılan denetimli yöntem olan ′belirsizliği sorgulama′. Bu yöntemde karar sınırına en yakın örnekler seçilir ve etiketlenmesi için insan uzmana sorulur. Biz çalışmamızda anlatılan bu iki aktif

öğrenme yöntemini de denetimsiz öğrenme için geliştirip REPERE verisi üstünde uyguladık.

Anlaşmazlığa gore sorgulama yönteminin denetimsiz öğrenme probleminde geliştirilmesinin ve problemlerinin çözümünün ilk adımı olarak iki adet denetimsiz kümeleme yöntemi eğitilmiştir. Kümeleme yöntemi olarak yığısal kümeleme(agglomerative clustering) ve K-medoid kümeleme yöntemleri kullanılmıştır.

Denetimli yöntemde anlaşmazlığa göre sorgulamada sınıflandırılar test aşamasında etiket atadığından anlaşmazlık direk anlaşılıp insan uzamana sorulurken, denetimsiz yöntemde etiketler bulunmamakta, etiketsiz kümeler bulunmakta. Bu durumda iki kümeleme yönteminin kümelerini karşılaştırabilmek için aynı etiketi temsil eden kümelerin birbiriyle işleşmesi gerektiği düşünülmüş ve eşleşme problem çözülmüştür. Kümelerin eşleştirme işlemi için yığınsal kümeleme yönteminden gelen her bir küme için K-medoid yönteminin çıktısı olan her bir küme ile ortak eleman sayısı bulunup toplam eleman sayısına bölünmüş ve kümeler arasında bir anlaşmazlık değeri matrisi oluşturulmuştur. Kümeler birlerine göre değerlendirildikten sonra oluşturulan benzerlik matrisi kullanılarak birbirine en uygun kümeler kararlı evlilik probleminin(stable marriage problem) çözüm algoritması uygulanmış ve en uygun küme çiftlerinin birbirine atanması garanti altına alınmıştır. İki kümeleme yönteminin ürettiği kümeler birbirlerine atanma işleminden sonra aralarında en fazla anlaşmazlık olan küme çifti içlerinden sorgu yapılması amaçlı seçilmiştir. Biz bu yöntemi ”en çok anlaşmazlığı sorgulama” (MDS) olarak adlandırdık. Bu aşamada hangi küme çifti içinden örnek seçileceği belirlendi ancak kümeler içinde hangi örneğin sorulacağını seçmedik.

Seçilen küme çifti içinden örnek sorgulanması için örneğin seçiminde belirsizliğe -uncertainty sampling- göre sorgulama yöntemini küme için uyguladık. Bu adımda seçilen küme çifti içinde ki örnekler tek tek taranarak diğer örneklere toplam uzaklığı en fazla olan örneği seçtik. Diğer örneklere toplam uzaklığın en fazla olduğu örnek bize küme içinde kümenin sınırına en yakın olan yani belirsizliği en fazla olan örneği verir. Deneyler sırasında görüldü ki küme içinde en belirsiz noktaları, yanı olasılığı en düşük noktaları seçmenin yanı sıra olasılığı en yüksek olan noktaları yani kümenin merkezine yakın olan, olasılığı en yüksek olan noktaları seçip sorgulamak da başarıyı arttırmakta. Bu nedenle belirsizliği en düşük olan noktalar seçilip sorgulama yapıldığı deneylerde bu çalışmada yapıldı.

Bu çalışmada önerilen yöntem [2] çalışmasında önerilen ”Big Cluster First” (BCS) yöntemiyle karıştırıldı ve iki yöntemin de güçlü olduğu bölümlerin olduğu görüldü. Bu nedenle BCS ve MDS yöntemleri birleştirilerek melez yöntemde önerilmiştir.

Deneylerde aktif öğrenme algoritmaları REPERE video verisi üstünde 4 farklı sınıflandırma problem üstünde yapılmıştır. Burda amaç videoda o anda kimin konuştuğunun düzgün etiketlenmesidir. Sınıflandırma problemlerinden ilki konuşmacı yani ses verisinden gelen öz nitelikler kullanılarak, videoda ki konuşmacının etiketlenmesi. İkincisi problemde ses verisinden gelen öz nitelikler kullanılarak videoda gözüken kişinin etiketlenmesi. üçüncü problemde görüntü verilerinden gelen öz nitelikler kullanılarak videoda gözüken kişinin etiketlenmesi ve son sınıflandırma probleminde görüntü verisinden gelen öz nitelikler kullanılarak o anda ki ses verisinin

etiketlenmesi. Anlaşılacağı üzere önerilen aktif öğrenme yöntemleri çok kipli problem üstünde uygulanmıştır.

Deney sonuçları 28 video için mikro F-measure, makro F-measure hesaplanarak değerlendirilmiştir. Ayrıca yöntemlerin başarıları kendi aralarında derecelendirilmiştir ve en iyi yöntem belirlenmeye çalışılmıştır. Deney sonucunda görülmüştür ki en başarılı yöntem ufak farkla da olsa kesinliği en yüksek örneklerin sorgulandığı melez yöntem olduğudur.

1. INTRODUCTION

In the last decade, video recording technologies have moved from analog to digital and video recording devices have proliferated among wide range of users. Another evolution is proliferation of high speed internet. Many regular internet users started to create videos after digital video evolution and they can share these videos over the internet. In addition, video streaming service providers, such as Youtube, personal blogs, news web sites, also contribute new data every day. According to the Youtube statistics, 100 hours of video are uploaded to Youtube every minute and over 6 billion hours of video are watched on Youtube every month [3]. The enormous amount of video data causes challenges in search and retrieval operations. In order to index and search videos with the required contents effectively, annotation process is applied on them. There are two main video annotation techniques; manual annotation, in which a human expert annotate a video manually and automated video annotation, in which various computational methods annotate a video. In manual annotation approach, an expert watches whole videos frame by frame and annotates topics, occurred events, persons, types (e.g music, news etc) and other informative data. On the other hand, computational methods may annotate many videos automatically without any or less help from human. Automated video annotation methods are faster but more error prone than the manual method. Even though manual approach is the most accurate one in terms of video annotation in many cases, annotating immense quantity of videos manually is a labor-intensive and time-consuming process.

Automatic annotation techniques such as Active Learning offers various solutions to overcome the excessive cost of manual annotation. Active learning is a semi-supervised machine learning technique that aims to reduce annotation/labelling costs. Conventional supervised or unsupervised machine learning methods use random annotation for learning phase. In active learning, an algorithm annotates videos automatically but asks for the labels of the most informative instances to human expert

in order to learn data distribution more accurate. By this way, experts may train learning model with less annotation effort.

Various active learning methods have been proposed to video annotation problem. One of the studies [4] applies active learning for video annotation by comparing uncertainty sampling, the most probable sampling and random sampling to video indexing. A study [5] proposes video retrieval and annotation system called LIGVID which uses two active learning methods: 'relevance sampling' and 'uncertainty sampling'. Another study integrates SVM based active learning for feature selection to solve the text classification problem [6]. However, active learning for feature selection fails in that study, because of the use of a wrong feature reduction technique called GainRatio Feature Selection [6].

In a study [7], active learning is applied on networked data, of which nodes are 'papers' and links are 'references to other papers'. It uses a method based on query by disagreement and reduces paper annotation costs for classifying research papers. A study [8] extends the traditional active learning framework by including feedback on features alongside labeling the instances. It focuses on the effects of feature selection and human feedback for features in the setting of text categorization and applies uncertainty sampling based methods.

This study is a sub-project of the project CAMOMILE [9], which targets to produce an annotation framework for 3M data where the letter 'M' stands for multimodal, multimedia and multilingual. In this study, we propose a cluster based unsupervised active learning approach as a selection strategy on REPERE [1] video dataset, which is created for the person identification problem in videos. Our study aims to identify who is speaking and who is on screen.

1.1 Thesis Organization

The rest of the thesis is organized as follows: Section 2 includes the notation for the study and a background of used dataset. Whereas,Section 3 includes related works on active learning and cluster matching, Section 4 introduces the methods. Section 5 includes the experiments and concludes the study.

2. NOTATION AND DATASET

2.1 Notation

Literally, the finite data set is $X = x_1, x_2, \ldots, x_n$ where the cardinality is $|X| = n$. The cluster set $C = \{\mathrm{C}_1, \mathrm{C}_2, \ldots, \mathrm{C}_k\}$ represents the cluster sets of the set X with the asumption $|\mathrm{C}_i| > 0$ for all $i = 1, 2, \ldots, k$. The set of all clustering of X is denoted as $B(X)$. C' is the second clustering for X where $C' = \{\mathrm{C'}_1, \mathrm{C'}_2, \ldots, \mathrm{C'}_\ell\} \in S(X)$. $\mathrm{M} = (mij$ is the confusion matrix of each clustering pair C, C'. The intersection between C_i and $\mathrm{C'}_j$ is a $kx\ell$ matrix where the ij^{th} element gives the number of the elements in the intersection of C_i and $\mathrm{C'}_j$.

$$m_{ij} = |\mathrm{C}_i \cap \mathrm{C'}_j|, 1 \leq i \leq k, 1 \leq j \leq \ell \tag{2.1}$$

The product of the clustering C_i and the clustering $\mathrm{C'}_j$ is denoted as $\mathrm{C}_i \times \mathrm{C'}_j$ and defines the coarsest common refinement of those two clusterings.

$$C \times C' = \{\mathrm{C}_i \cap \mathrm{C'}_j | \mathrm{C}_i \in C, \mathrm{C'}_j \in C', \mathrm{C}_i \cap \mathrm{C'}_j \neq \emptyset\} \tag{2.2}$$

2.2 Dataset

We use video dataset from the REPERE challenge, which aims to provide a benefit to research on person identification in videos for multimodal conditions. The REPERE challenge tries to find answers to questions “Who is speaking?” “Who is present in the video?” and etc. by use of various information on speech and image modals. The REPERE Corpus occurs from 28 videos, which include 7 different types of shows such as news, talk show etc. and various numbers of participants from 3 persons to dozens in a video. In addition, the length of the videos has a range from 3 to 30 minutes,

Figure 2.1: Example frames for face segmentation and OCR [1].

which naturally causes various numbers of annotations for each video approximately 20 to 100 frames.

As distinct from other learning approaches, active learning interactively asks for annotation from an expert depending upon a selected strategy. So, an initial set of annotated data is needed for these queries. In general, embedded texts in videos can point the name of a represented speaker. However, it is expensive and time-consuming to obtain text annotations manually. Therefore, we use the extracted overlaid texts in videos by an Optical Character Recognition (OCR) system [10].

The used text detection method is adopted from the study [11]. It includes a Sobel filter and erosion/dilation operations that are employed for coarse detection following by filtering out false positive text boxes. Moreover, Google OCR system called Tesseract is used for text recognition. After the OCR system detects and extract texts automatically, human annotators verify and improve the results.

Speech and face segmentations processes are applied on videos to gather feature set for both training and testing. For Speech segmentation, speech tracks are gathered by splitting signals into acoustically homogeneous segments. For each video, the similarity matrix between speech tracks is calculated by using BIC criterion [12] with single full covariance Gaussians. The similarities are normalised into the interval [0,1].

For face segmentation, the first and the fifth frames of each shot is scanned to find a face of which initial pose can be frontal, half-profile or profile. Moreover, the segmentation approach uses online tracking, which determines the current position and the location

of a face in a frame by using the information of previous frame. After gathering a face image, 9-point mesh (an eye: 2 points, a nose: 3 points and lips: 2 points) is imposed on it. Then, a 490 dimensional feature is calculated with HOG descriptor. An average descriptor is calculated by using found features of a face on the same shot. For each face track, the number of feature dimension is decreased from 490 to 200 with the help of LDML approach [13]. At last, the similarity matrix between face tracks is calculated with Euclidean distance. The similarities are normalized into the interval [0,1].

The third similarity matrix occurs from correlation scores between the faces and the speakers in the interval [0,1] to build a multimodal clustering. Too small faces are eliminated for this distance matrix. The change in the color histograms of the region of a speaker's lips indicates person talking. Moreover, the size of a face and proximity to the center of the screen give us the clue that the current speech track is associated to the current face track.

The test set of REPERE challenge includes 1229 annotated frames from three hours videos. Instead of annotating all videos frame by frame, every 10 seconds a frame is selected and annotated by human annotators.

Literally, each instance of the dataset $X = x_1, x_2, \ldots, x_n$ represents a face track and $Y = y_1, y_2, \ldots, y_n$represents a speech track. An instance x_i has 200 dimensional feature vector $x_i = (x_i^1, x_i^2, \ldots, x_i^{200})^T$ Likewise an instance y_j has feature vector $y_i = (y_i^1, y_i^2, \ldots, y_i^m)^T$. The distance matrices D_f, D_s and D_{fs} represents the distances, respectively, between face-to-face, speech-to-speech and face-to-speech tracks in a video. The size of the distance matrices and speaker track number ST, face track number ST and overlaid name numbers ON are given in the Table 2.1.

Table 2.1: The video list and the size of distance matrices which are inputs for our algorithms.

Video Name	D_f	D_s	D_{fs}	FT	ST	ON
BFMTV BFMStory 2012-01-10	483×483	348×348	1020	759	372	36
BFMTV BFMStory 2012-01-23	507×507	329×329	1037	782	338	34
BFMTV BFMStory 2012-02-14	498×498	342×342	1024	753	348	33
BFMTV BFMStory 2012-02-20	566×566	330×330	1104	871	341	30
BFMTV CultureEtVous 2012-01-13	40×40	36×36	94	69	38	3
BFMTV CultureEtVous 2012-01-16	92×92	39×39	178	163	42	5
BFMTV CultureEtVous 2012-01-17	30×30	44×44	70	48	49	5
BFMTV CultureEtVous 2012-01-18	50×50	45×45	128	102	47	6
BFMTV CultureEtVous 2012-01-19	49×49	38×38	101	4	78	41
BFMTV CultureEtVous 2012-02-14	59×59	48×48	136	9	104	48
BFMTV CultureEtVous 2012-02-15	94×94	34×34	183	4	180	34
LCP CaVousRegarde 2011-12-20	352×352	327×327	681	397	332	30
LCP CaVousRegarde 2012-01-19	306×306	289×289	638	371	290	22
LCP CaVousRegarde 2012-01-25	356×356	350×350	747	428	352	25
LCP EntreLesLignes 2011-12-16	428×428	251×251	788	560	253	15
LCP EntreLesLignes 2012-01-27	417×417	264×264	840	599	267	14
LCP EntreLesLignes 2012-05-11	405×405	300×300	949	960	305	13
LCP LCPInfo13h30 2012-01-24	159×159	158×158	337	212	163	13
LCP LCPInfo13h30 2012-01-25	262×262	168×168	460	331	172	10
LCP LCPInfo13h30 2012-01-27	225×225	170×170	456	306	177	11
LCP PileEtFace 2011-11-19	207×207	210×210	441	243	211	40
LCP PileEtFace 2011-12-01	293×293	254×254	587	389	255	51
LCP PileEtFace 2012-01-12	277×277	268×268	580	330	269	52
LCP PileEtFace 2012-01-19	281×281	244×244	518	302	246	43
LCP PileEtFace 2012-01-26	283×283	270×270	624	380	273	50
LCP TopQuestions 2012-01-25	258×258	130×130	480	364	134	11
LCP TopQuestions 2012-02-14	172×172	153×153	395	249	155	13
LCP TopQuestions 2012-02-22	214×214	155×155	439	308	158	15

3. RELATED WORK

In this chapter, we review some fundamental concepts for this study. We give a brief overview of active learning, followed by a general discussion on active learning scenarios and basic active learning strategies.

3.1 Active Learning

Active learning mainly discovers the most informative instances in terms of data distribution in data, while training a model in an iterative manner to deal with the cost of labeling. Determining the most informative instances and querying human annotators iteratively decreases the needed number of labeled data. Moreover, it helps the learner algorithm for learning the underneath distribution relatively faster than supervised methods. A nice example [14] about the human colonial, which arrives a solar planet, explains the semantic behind the active learning. In this scenario, the planet is habitable and includes a large amount of eatable vegetation. Important amount of the food comes from a plant of which fruits are smooth, round and irregular. The smooth and round fruits are delicious and good for humans. On the other hand, irregular fruits cause sickness. It is very indispensable to classify fruits as safe and noxious for the favor of the colony.

According to the PAC learning framework [15], if data distribution can be perfectly classified by some hypothesis function h in the hypothesis function set H then it is enough to test $\mathcal{O}(1/\varepsilon)$ randomly selected instances, where ε is the maximum desired error rate. In other words, hundreds of fruits might be tested to be able to achieve 99% accuracy for 'fruit safety' classifier. Unfortunately, this experiment can cause lots of poisoned people. Conversely, instead of choosing the fruits randomly for specifying fruit irregularity threshold θ, we can find and use the most informative fruits.

All things considered, we are able to formulate the classification problem of 'safe' and 'noxious' food by using simple binary classifier. Literally, each instance of the dataset

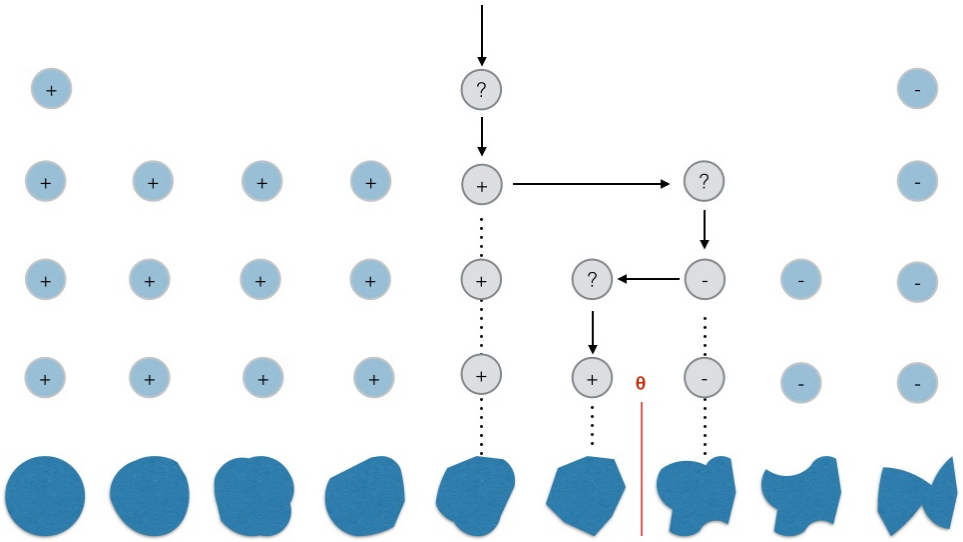


Figure 3.1: Detecting safe shaped food threshold θ with binary search.

$X = x_1, x_2, \ldots, x_n$ represents a harvested fruit. Each data instance x_m has a hidden label $y \in \{1(safe), 0(noxious)\}$. Thus, the binary classification problem is defined as the set of x_m, y_m pairs with the function mapping $h : X \rightarrow Y$, parameterized by a threshold θ, which represents the best threshold to be able to decide 'safe' and 'noxious' fruits. Supervised learning based methods need more human subjects to specify the threshold. Fortunately, binary search reduces the threshold search cost in this problem as seen in Figure 3.1.

The study [14] indicates that binary search speeds up the finding errorless classifier for a number of $\mathcal{O}(\log_2 \frac{1}{\varepsilon})$ training instances where ε indicates the error rate. By this way, there is no need to use all of the training instances. For example, previous supervised methods use 100 human subjects to taste 100 fruits for obtaining 99% accuracy. On the other hand, the same accuracy can be achieved with 6 or 7 human subjects by using active learning based classifiers. In the end, using the most informative fruits enables a substantial reduction in the number of sick people.

3.1.1 Active Learning Scenarios

Active Learning scenarios defines how to ask a selected instance to a human expert. Mainly, there are 3 different Active Learning scenarios namely query synthesis, stream based selective sampling and pool based sampling which explained in the following sub sections in detail. In this study, we use stream based selective sampling.

3.1.1.1 Query Synthesis

D. Angluin's 'Query and concept learning' [16] article describes the query synthesis. This scenario investigates learning with membership queries. In query synthesis learning scenario, learner may request label membership for any labeled data instance in the input space including queries that the learner starts from the beginning [17].

Query synthesis is good if learner has an exact definition of the input space. I realise in my literature search that query synthesis is often used. Especially recent works rarely use query synthesis. [14] gives a good example about query synthesis. [14] applies membership query learning with human oracles to train a neural network classifier of handwritten characters. The problem is many of query images generated by the learner can not recognised by oracle. They encountered an unexpected problem: many of the query images generated by the learner contains no recognisable symbols; these are artificial hybrid characters with little or no natural semantic meaning for oracle. For example; hand writing tools can express 5, 8, or 9 instead of the real image of a number. It stands to reason that undetermined images could help the learner discriminate among the different characters, if people were able to discriminate among them as well. Similarly, text or speech recognition tasks are not suitable for query synthesis because of same sample generation problem.

3.1.1.2 Stream Based Selective Sampling

Other query scenario is stream based selective sampling. Stream based selective algorithm is easy to understand. Learning starts with random labeled data, when the learning process, learner decides to ask one sample to oracle. The key assumption is that obtaining an unlabeled instance cost is none (or inexpensive), so it can be sampled from the actual distribution, and then the learner can decide whether or not to request its label. If the input distribution is uniform, selective sampling cannot offer any advantage over query synthesis. But if the input distribution is non-uniform, selective sampling gives better result than query synthesis [14].

3.1.1.3 Pool-Based Sampling

Stream based selective sampling is needs more computation and time. Real world problems are not suitable stream based selective sampling. In real world problems, starting with small unlabeled data to training then learner ask oracle to large collections of unlabeled data. Learner creates a query pool when training [14].

The study [14] explains difference between stream based selective sampling and pool based sampling difference in his book. Main difference between stream based and pool based active learning is selection of query sample. Stream based query ask one sample at a time. But pool based method asks set of sample to expert after each learning iteration. Stream based query method looks more precise then pool-based because they ask one sample after each iteration this is why asked sample quality is guaranteed. But pool based sample can ask some noise or meaningless samples. On the other hand, stream based querying need more training time, because each annotation followed by training phase. This is why pool based querying is more popular if annotation cost is not too high unlike video annotation.

3.1.2 Sampling Strategies

We explain how to ask scenarios in section two. Most important part of active learning is how to choose or decide which samples are best for the asking to oracle. I will explain three strategies, uncertainty sampling, query by disagreement, query by committee.

3.1.2.1 Uncertainty Sampling

Uncertainty sampling is proposed by [18]. This algorithm queried most confident samples for learner. Main idea is finding most informative instances from unlabeled input. So we can find most informative instances near the decision boundaries. If learner chooses instances near of decision boundary, probability near 0.5 for two class problem near decision boundary, learner can ask these confused points to oracle for labeling, so learner can train better in next iteration in theory.

The book [14] refers three measures of uncertainty in his active learning book.

- **Least Confident**

  Least confident method is a basic uncertainty sampling strategy to query the instance whose predicted output is the least confident:

$$x^*_{CL} = argmin_x P_\theta(\hat{y}|x)$$

$$= argmax_x 1\text{-}P_\theta(\hat{y}|x) \ (1).$$

  With the highest posterior probability under the model *j*. In other words, this strategy prefers the instance whose most likely labeling is actually the least likely among the unlabeled instances available for querying [14].

- **Margin**

  A different active learning strategy is based on the output margin:

$$x^*_M = argmin_x [P_\theta(\hat{y_1}|x) - P_\theta(\hat{y_2}|x)] \ (2)$$

  where $\hat{y_1}$ and $\hat{y_2}$ are the first and second most likely predictions under the model, respectively Margin sampling addresses a shortcoming of the least confident strategy by incorporating the second best labeling in its assessment [14].

- **Entropy**

  Perhaps the most general (and most common) uncertainty sampling strategy uses entropy [19], usually denoted by H, as the utility measure:

$$x^*_H = argmax_x H_\theta(Y|x)$$

$$x^*_H = argmax_x \text{ - } \sum_y P_\theta(y|x) \, log P_\theta(y|x) \ (3)$$

  Where "y" ranges over all possible labels of "x". Entropy is a measure of a variable's average information content. As such, it is often thought of as an uncertainty or impurity measure in machine learning.

3.1.2.2 Query by Disagreement

Query by disagreement (QBD) proposed by [20]. Disagreement assumes the stream-based selective sampling scenarios so unlabeled data streams to oracle for labeling. Main idea is setting up two hypotheses; one hypothesis is most general hypothesis on the other hand other hypothesis sets up most specific hypothesis. Two classifiers work together but if one of classifier is disagree with other classifier instance is asked to oracle for labeling.

The study [14] indicates, QBD needs perfect approximate to boundaries for measure disagreement among all hypotheses in the version space, even if we imperfectly approximate it with two extreme hypotesis, most general and most specific QBD fall in trouble. Query by committee relaxes this assumption using multi classifiers.

3.1.2.3 Query by Committee

More theoretically-motivated query selection framework is the query by committee (QBC) algorithm [21]. Query by committee approach foundation is set of different classifiers. In this approach, set of different classifiers train for the same data set. In labelling section all classifiers propose own label to same sample. Same sample label agreed by all or most of classifiers but some sample cause conflict between classifiers. This disagreed samples are most informative data. Because those samples are near decision boundary in theory and we know that decision boundary is the most informative section of hypothesis space [14]. Bagging like method works well in query by committee [22].

Disagreement of classifiers need a measure method. Two dominant methods used for the measuring disagreement in classification task. First one is based on entropy, which called vote entropy. This measurement is most used measurement in literature. Formulation is given [14]:

$$x^*_{VE} = argmax_x \text{-} \sum_y \frac{vote_c(y,x)}{|C|} log \frac{vote_c(y,x)}{|C|}$$

$$vote_c(y,x) = \sum_{\theta \varepsilon C} 1_{h_\theta(x)=y} \quad (4)$$

In this formulation, y is all possible labels, |C| size of the committee.

Another disagreement measure is Kullback-Leiber (KL) [23] . This method tests measurement of the difference between two probability distributions. In this case , KL quantify disagreement as the average divergence of each committee member θ prediction from that of the consensus C. Also formulation is given as the following;

$$x^*_{KL}=\underset{x}{argmax}\frac{1}{|C|}\sum_{\theta \epsilon C} KL(P_\theta(Y|x)\,||P_C(Y|x))$$

$$KL(P_\theta(Y|x)\,||P_C(Y|x)) = \sum_y P_\theta(y|x) log\,\frac{P_\theta(y|x)}{P_C(y|x)} \quad (5)$$

[14] mentions about KL that some cases probability distributions is not uniform even if not consensus output may be uniform. This is difference between vote entropy and KL divergence.

3.2 Cluster Matching Metrics

[24] proposes theoretic "Variation Of Information"(VI) criterion for comparison of two clustering on the same data set. Proposed VI criterion measures and compares the qualities of two different C and C' clusters to determine the better clustering method in terms of information gain and loss. In addition, the study gives a nice literature review about the comparison methods of previous studies and groups them. The first group of comparison methods uses Counting Pairs of compared cluster methods. However, Counting Pair methods do not give better results than other comparison metrics because of the asymmetry. Pair of points from X described below for all cases:

- S_{11} demonstrates the number of the data point pairs that are in the same cluster under both C and C'
- S_{00} demonstrates the number of the data point pairs that are in the different clusters under both C and C'
- S_{10} demonstrates the number of the data point pairs in the same cluster under C, but in different clusters under C'
- S_{01} demonstrates the number of the data point pairs in the same cluster under C', but in different clusters under C

The sum of all the cases always must be;

$$S_{11} + S_{00} + S_{10} + S_{01} = n(n-1)/2 \tag{3.1}$$

By considering the Eq.**3.1**, [25] proposes two asymmetric comparison criteria Eq.**3.2** and Eq.**3.3** which gives the probability of point pairs in the same cluster under both C and C'.

$$W_I(C,C') = \frac{S_{11}}{\sum_i n_i(n_i-1)/2} \tag{3.2}$$

$$W_I I(C,C') = \frac{S_{11}}{\sum_j n'_j(n'_j-1)/2} \tag{3.3}$$

[26] proposes a comparison metric called Rand Index which compares the result of a classification scheme with a correct classification (See Eq.**3.4**). Rand Index depends the number of clusters and the number of data points. In the equation, the function R gets values between [0,1] where $R = 0$ indicates that the two clustering algorithms are identical.

$$R(C,C') = \frac{2(n_{11} + n00}{n(n-1)} \tag{3.4}$$

[24] mentions another criteria called Jaccard Index as in Eq.**3.5**. Jaccard Index gets values between [0,1] where $J = 0$ indicates that the two clustering algorithms are identical.

$$J(C,C') = \frac{S_{11}}{S_{11} + S_{01} + S10} \tag{3.5}$$

[27] improves the Jaccard Index as in Eq.**3.6**

$$M(C,C') = \sum_i {n_i}^2 + \sum_j {n'_j}^2 - 2\sum_i \sum_j {m_{ij}}^2 \tag{3.6}$$

[24] mentions Set Matching Criterion family as the second group of comparison methods which compare clusters using cluster sets not data points. The first Set Matching Criterion called Larsen and Aones criterion which is asymmetric

(SeeEq.**3.7**). The criterion function L gets values between [0,1] where $L = 1$ indicates that the two clustering algorithms are identical.

$$L(C,C') = \frac{1}{k}\sum_{i} max_j \frac{2mij}{n_i + n'_j} \tag{3.7}$$

[28] proposes criterion in the Eq.**3.8** which finds the best matches of each cluster C in C' and is symmetric. It scans element of m_{ij} of the contingency table in decreasing order. Largest elements of m_{ij} called as nab. All row and column matches are listed until to reach $min(k,\ell)$. $match(i)$ function is index of C'_i in C' that matches cluster C_i. The criterion function H gets values between [0,1] where $H = 1$ indicates that the two clustering algorithms are identical.

$$H(C,C') = \frac{1}{k}\sum_{j=match(i)} m_{ij} \tag{3.8}$$

Another symmetric criterion introduced by [29] is given in Eq.**3.9**. $D = 0$ indicates that the two clustering algorithms are identical, otherwise D value is smaller than $2n$.

$$D(C,C') = 2n - \sum_{i}^{k} max_j m_{ij} - \sum_{j}^{\ell} max_i m_{ij} \tag{3.9}$$

All the three matching symmetric method functions L in Eq.**3.7** , H in Eq.**3.8** and D in Eq.**3.9** finds the best matching sets first for each cluster and compares the clusters using best matching values. The study [24] mentions the problem of these three criterion do not compute "unmatched" part of each cluster. So these metrics do not know what happens clusters unmatched points. It introduces VI criterion for solving asymmetric and unmatching problem.

Variation of information is aimed to compare clusters C and C' using disagreement. It uses entropy based method for measuring disagreement. First step of finding entropy of cluster is calculate probability of cluster sets:

$$P(i) = \frac{n_i}{n} \tag{3.10}$$

Thus, we can calculate the entropy of each cluster as;

$$H(C) = \sum_{1}^{k} P(i) log P(i) \tag{3.11}$$

The formula in Eq.**3.11** shows the entropy measured in bits and entropy does not depend of number of points. this calculation is basement to calculation of mutual information between two clustering.

Calculating of mutual information need $P(i, j)$ this is joint distribution of the random variables.

$$P(i, j) = \frac{|C_i \cap C'_j|}{n} \tag{3.12}$$

Mutual information of associated random variables described in below;

$$P(i, j) = \sum_{i=1}^{k} \sum_{j=1}^{\ell} P(i, j) log \frac{P(i, j)}{P(i)P'(j)} \tag{3.13}$$

[24] introduces variation of information using mutual information and entropy associated with cluster C.

$$VI(C, C') = [H(C) - I(C, C')] + [H(C') - I(C, C')] \tag{3.14}$$

We can see that I(C,C') is intersection of H(C) and H(C') this is why [24] proposed criterion find disagreement of two cluster subtracting mutual information from each entropy associated cluster. It examined VI under 12 property like non-negativity, symmetry, triangle inequality, n-invariance, upper bound of VI, splitting a cluster, collinearity of join, linearity of composition.

Another study about evaluation metrics to compare clustering algorithms done by [30]. Amigo et al. defines four new formal measures for clustering comparison, writers also examines some comparison metrics using newly defined constraints. Amigo et al. also explain why those metrics can not satisfy all constraints in first part of study. Second part of study Authors analyze BCube algorithm because BCube satisfies all of four constraints also BCube can compare clustering that has overlapping problem.

[30] tested some comparison metrics - explained before- using 12 mathematical constraints. It defines higher level formal constraints than [24] constrains because authors expect some rules for constraints:

A Constraint should be intuitive and tests clearly limitation of each metric. So system developer can be sure which constraints are important for the specified task. Constraint should be prove formally. Metric families must be discriminated by constraints. This property is useful for system developer because system developer can choose right metric if previous tested family does not work properly for the problem.

Using this point of view, authors define those formal constraints on evaluation metrics for clusterings.

Constraint one is cluster homogeneity: This constraint is purpose of clustering algorithms, clusters must be homogeneous. In a word, all clusters have same item. Second constraint is cluster completeness: First constraint measure homogeneity but same class items can be clustered in two or more cluster. In many case, big one cluster is better than two or more clusters for same class items. Third constraint is rag bag . Rag bag is clustering noise points to same cluster that means rag bag is set of 'others', 'unclassified', 'miscellaneous' categories. Constraint four is cluster size versus quantity. Big cluster and small error should be preferable to large number of small error in small clusters. For example if two clustering algorithms has same error number for same data, clustering algorithm that has bigger cluster is better than other algorithm.

Authors examine families of metrics under those constraints, results:

Set matching metric families can not satisfy constraints 2 and 3. Constraint 1 is satisfied by purity measure. Counting pair based metrics family-like Rand, Jaccard Coefficient, Folkes and Mallows can satisfy constraints 1 and 2 but they do not satisfy 3 and 4. Entropy based metric family like Meila's Variation of information fail to satisfy constraints 2 and 4. Evaluation metrics family based on edit distance can not satisfy constraints 1 and 3. [30] introduce BCubed metrics - metric that is mixed family of metrics- can satisfy all of four constraints.

4. METHODOLOGY

Supervised query by disagreement method (QBD) [14] is one of the active learning methods in the literature. QBD uses two different classifiers and asks the label of a disagreed point to an expert. We introduce unsupervised query by disagreement method to accelerate the learning phase of the automatic video annotation by using unsupervised learners. When it is used with supervised learners, Query by Disagreement (QBD) lets the learners label each instance and compares the learners' outputs for the same instance to detect disagreement instances. However, since unsupervised methods do not generate labels for each data instance as output of a learner, using QDB on unsupervised learners is a new and a challenging problem. We propose a novel approach for QBD on unsupervised learners and apply the solution on multimodal video data.

In our study, we use the following steps during an Active Learning Cycle: clustering, cluster matching, disagreement measurement between clusters, selection of the most disagreed clusters, selection of an instance to be queried from the selected cluster. The arhitecture of our system is depicted in Figure 4.1

4.1 Clustering Algorithms

Active learning cycle begins with the clustering of the data. We use two clustering algorithms: Agglomerative Clustering and K-Medoid Clustering. Agglomerative clustering is a hierarchical clustering algorithm and works with bottom-up approach. Therefore, it does not need to take a fixed number of clusters as input. For a given distance threshold, Agglomerative Clustering estimates the number of clusters, which we use as the value of the K-Medoid Clustering's number of clusters parameter k.

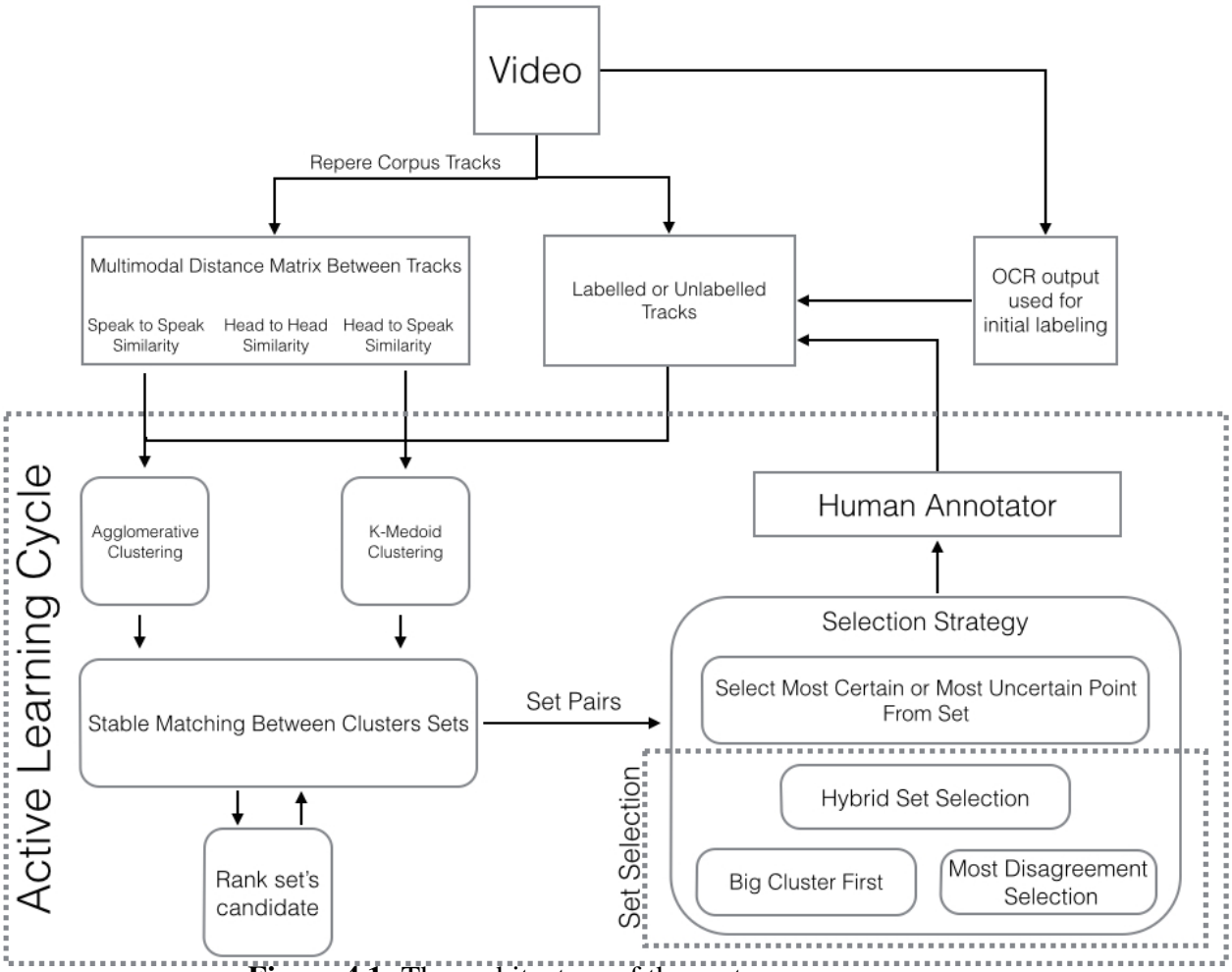


Figure 4.1: The architecture of the system.

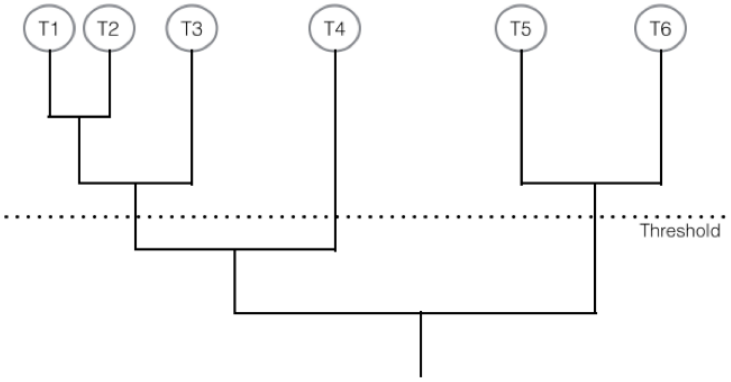


Figure 4.2: The illustration of Agglomerative Clustering.

4.1.1 Agglomerative Clustering

Agglomerative Clustering method clusters instances up until reaching a given distance threshold and returns the number of clusters. In our dataset, each video data has variable size of classes. However, the other clustering algorithm that we use, K-Medoid, needs number of clusters as input. Thus, we let the Agglomerative Clustering finds the number of the clusters and use it as input in K-Medoid Clustering. The generic Agglomerative Clustering algorithm is given in the Figure 4.3

```
Initialize
Assign each data point as a cluster
while until reaching a given distance threshold do
   Find the most similar pair of clusters
   Merge them into a single cluster
   Compute new similarities of the new clusters
end while
```

Figure 4.3: Pseudo-code of Agglomerative Clustering.

4.1.2 K-Medoid Clustering

K-Medoid Clustering algorithm needs an initial k parameter which corresponds to the number of clusters. However, our proposed unsupervised query by disagreement needs at least one clustering method that does not need to know the number of clusters initially. Therefore, we use agglomerative clustering to gather the number clusters and use it as K-Medoid's parameter k.

```
Initialize
k points are randomly selected in n instances as medoids
while Is medoids shifted  do
   Each data point is associated to the closest medoids
   for each medoid m  do
      for each datapoint 'o' is not a medoid do
         swap m and o and calculate the total costs of the configuration
      end for
   end for
end while
```

Figure 4.4: Pseudo-code of K-Medoid algorithm.

In order to find the clustering disagreements between the algorithms, we first solve its complementarity problem by finding the agreed instances on correspondent clusters the of algorithms. We apply a cluster matching algorithm to measure the similarities between the clusterings produced by the Agglomerative and K-Medoid algorithms. We find the intersection of clusters by calculating the cluster similarity metric given in Equation **3.8**. For our dataset, the cluster pairs' similarity matrix using best match measurement is given in Table 4.1.

Table 4.1: A sample cluster pair similarity matrix for agglomerative and K-Medoid algorithms.

		K-Medoid		
		Cluster 1	**Cluster 2**	**Cluster 3**
Agglomerative	**Cluster 1**	0.437	0.094	0.321
	Cluster 2	0.318	0.432	0.176
	Cluster 3	0.134	0.126	0.173
	Cluster 4	0.320	0.284	0.421

After finding the cluster pairs' matching similarities, we need to couple the pairs using cluster pair similarity matrix that we constructed using 'Best Match'. We calculate the similarity matrix between the sets of two clustering algorithms but we do not assign Agglomerative Clustering's sets to K-Medoid Clustering's sets one by one yet. The 'assigning a set to another set' problem perfectly fits into 'Stable Marriage Problem'. In real life, finding the most coherent man and woman pair is the key of a stable

marriage. However, n men and n women stable marriage is an optimization problem because all candidates are assigned to the most coherent candidates. Each man and woman ranks all of the opposite members from 1 to n and a stable marriage occurs when a man and a woman could not find any other better partners in opposite sex within $n-1$ candidate in terms of rankings.

Literally, we need to find $j = match(i)$ which denotes the cluster C'_j to which the C_i should be matched. Then, in order to assign clusters between Agglomerative and K-Medoid clusterings, we adopt and apply the Gale-Shapley algorithm ([31] , [32]) which solves the 'Stable Marriage Problem'. Gale-Shapley algorithm guarantees that the solution obtained is perfect (i.e. everyone gets married) and stable.

Let the clustering produced by the Agglomerative Clustering be $C = C_1, C_2, \ldots, C_k$ and the Clustering produced by the K-Medoid Clustering be $C' = C'_1, C'_2, \ldots, C'_k$. Gale-Shapley algorithm requires that each cluster C_i ranks the clusters C'_j and vice versa. We use the cluster confusion matrix entries (Equation **2.2**) in order to produce these rankings and then apply the Gale-Shapley algorithm to produce a cluster matching. The preferences of each Agglomerative cluster over K-Medoid clusters are denoted as P. The preference of the cluster C_1 may be shown as $P(C_1) = C'_2, C'_4, C'_1, \ldots$ where each cluster in the list $P(C_1)$ comes from the opposite cluster set C' and in descending order in terms of the candidates' scores. In the same manner, the preferences of each K-Medoid cluster over agglomerative cluster set C are listed. The preference list of the whole agglomerative cluster set is denoted as $P_a = P(C_1), P(C_2), \ldots, P(C_i)$. Likewise, the preference list of the whole K-Medoid cluster set is denoted as P_k. The output of the Gale-Shapley is a matching $M = [m_1, m_2, \ldots, m_k]$ where, $m_i \in 1, \ldots, k$, $m_i = j'$ if C_i is matched with C'_j in the stable matching. Stable matching algorithm for two clusters C and C' is given in Figure 4.5.

4.2 Query Instance Selection Strategies

In order be able to select which instance to query, we first select the most informative cluster of the cluster pair, and then we select the most informative instance in the selected cluster.

```
for each cluster C_i ∈ C do
   P(C) ⇐ BestMatch(C_i,C')
   Add P(C) into P_a
end for
for each cluster C'_i ∈ C' do
   P(C') ⇐ BestMatch(C'_i,C)
   Add P(C') into P_k
end for
while unstable do
   unstable ⇐ false
   for each cluster C_i ∈ C do
      if C_i has a pair cluster in M or P(C_i) = ∅ then
         continue
      end if
      unstable ⇐ true
      C'_i ⇐ the first element of P(C_i)
      if C_i has not a pair cluster in M then
         m_x ⇐< C_i, C'_j >
         Add m_x into M
         m_y ⇐< C'_j, C_i >
         Add m_y into M
      else
         m ⇐ current match for C'_i
         if Rank of < C'_j, C_i > > Rank of m then
            m_x ⇐< C_i, C'_j >
            Add m_x into M
            m_y ⇐< C'_j, C_i >
            Add m_y into M
            Delete current match m from M
         end if
      end if
   end for
end while
```

Figure 4.5: The stable matching algorithm.

4.2.1 Cluster Selection

We score clusters using different methods and select an instance from the cluster with the highest score.

4.2.1.1 Big Cluster First

The study [2] proposes 'Big Cluster First' (BCS) selection strategy which calculates a score using the size of a set and the number of annotated instances in that set. The method selects an instance from a minimum scored set by asking human expert. The BCS strategy score for a cluster C_i is calculated as shown in Equation **4.1**.

$$BCS(C_i) = \frac{NumberOfAnnotation(C_i)}{Size(C_i)} \tag{4.1}$$

4.2.1.2 The Most Disagree Selection

In theory, the most disagreed cluster pair gives us the most uncertain points because two stable matched clusters have lots of disagreed instances. MDS method chooses the cluster C_i which has the highest DS score. For measuring the disagreement between a pair of matched clusters C_i, C'_j where $j = m_i$, we calculate the disagreement score (DS), disagreement score list (DSL) and most disagree pair (MDS) as the following;

$$m_x = < C_i, C'_j >, \forall m \in M \tag{4.2}$$

$$DS(< C_i, C'_j >) = \frac{C_i \cap C'_j}{C_i \cup C'_j} \tag{4.3}$$

$$DSL\left\{DS(< C_i, C'_j >) | \forall m \in M : m_x = < C_i, C'_j >\right\} \tag{4.4}$$

$$MDS = (< C_i, C'_j >), where DS_s(< C_i, C'_j >) = min(DSL) \tag{4.5}$$

where

$$DS_s(< C_i, C'_j >) = min(DSL) \tag{4.6}$$

4.2.1.3 Hybrid Cluster Selection

BCS method performs better during the initial stages of active learning where the number of labeled instances are very few and labeling instances on big clusters help label a lot of instances. On the other hand, MDS performs better when clusters contain more known instances, because cluster disagreement has a correlation with the label assignment. Therefore, we introduce two hybrid cluster selection methods: Soft Hybrid Selection (SH) and Hard Hybrid Selection (HH).

In order to combine or compare the different BCS and MDS clustering scores, we apply the z-score normalization on them.

We know *BCS* performs better than random selection for early iterations in active learning cycle. On the other hand, *MDS* is not good at early stages but *MDS* performs good after few iterations later. Because of this manner, we develop the **Soft Hybrid Selection** method. Contracting SH method needs *BCS* like method for cluster pairs. This is why we propose big cluster pair first selection method for use in early iterations. The Big Cluster Pair First method is the derivative of *BCS* for paired clusters. The *BCPS* method calculates a score for each pair and selects lowest scored pair.

A cluster pair score is depicted as $S_p = A_p/N_p$ where A_p is the total number of annotation in a cluster pair and N_p is the total number of instances in a paired cluster. Soft hybrid method combines *BCPS* and *MDS* scores using weight. However, *MDS* and *BCPS* scores' mean and standard deviations are not comparable since they use different scales.Therefore, z-score normalization is applied over both *BCPS* and *MDS* scores for solving the scale problem. After that we calculate SH score (SHS) for each matched pair score using the given formula **4.7**. Normalized *BCP* and *MDS* scores are denoted as S_n and DS_n, respectively. The soft hybrid score SHS is calculated as;

$$SHS = ((1-\alpha) \times S_n) + (\alpha \times DS_n), 0 < \alpha < 1 \tag{4.7}$$

This formula clearly shows that if the value of α is close to 1, selection shifts to *MDS* while lower α value shifts SH selection to *BCPS*. α value must be close to zero at early stage of active learning for selecting cluster pair from *BCPS*. Calculation of α value is dependent to number of the iterations.

The **Hard Hybrid Selection** method uses *BCS* method for the given number of iteration. Important part is *BCS* uses only agglomerative clustering clusters. When active learning strategy cycle reaches to switching point, algorithm starts to use *MDS* for selection. In other words, hard hybrid uses *BCS* for early stages and switches to *MDS* later.

4.2.2 Instance Selecting Strategy

Selecting the most informative clusters with the defined cluster selection methods is the first step of the selection strategy. The next step is determining the most informative instances in the most informative clusters. The instances from the center of a cluster are more 'certain' than the instances close to the cluster boundary in terms of class knowledge. However, since the entropies of uncertain instances are relatively higher than the entropies of certain instances [14], the most informative instances are actually on the region around the boundary of a cluster. Therefore, for each instance in a cluster, we add the distance to other instances in the same cluster and choose the instance with the highest sum, which gives us the most 'uncertain' instance, the instance which is the farthest away from all the other instances. Instance selection strategy can select the most 'uncertain' or the most 'certain' instance from a cluster or a cluster pair.

5. EXPERIMENTS AND RESULTS

We evaluate the performance of selection strategies over active learning steps on REPERE Corpus data.

5.1 Evaluation Metrics

The confusion matrix also known as contingency matrix is basic metric of measurement of classifiers performance. The confusion matrix is an error for visualisation of classifier performance. The confusion matrix has class size of rows and columns. Each row contains total number of actual class value. Each column gives number of classifiers prediction. The confusion matrix reports four different data, $truepositives$, $falsepositives$, $truenegatives$, $truenegatives$. True positive indicates both actual class and the predicted class are positive. On the contrary, true negative indicates both actual class and predicted class are negative. On the other hand, false positive occurs when the actual class is negative but predicted class is positive. Likewise, false negative indicates the actual class is positive but predicted class is negative.

In this study, we use F-measure, which is an indicator of test set accuracy and harmonic mean of precision p and recall r values, to evaluate the algorithms. Precision shows how much relevant instances are retrieved among all retrieved instances. Regardless, Recall shows how much retrieved instances are relevant among all relevant instances. F-measure is calculated as in the Equation **5.1**

$$F - Measure = 2 \times \frac{Precision \times Recall}{Precision + Recall} \tag{5.1}$$

where

$$Precision = \frac{TP}{TP+FP} \tag{5.2}$$

and

$$Precision = \frac{TP}{TP+FN} \tag{5.3}$$

5.2 Results

We have evaluated the performance of selection strategies over active learning cycles using multimodality on 28 videos from 7 different TV programs. Each video has three different similarity matrices namely face-to-face, speech-to-speech and face-to-speech. We have run four experiments with regard to matrices in this order: (1) face score for face track annotation (FF), (2) face score for speaker track annotation (FS), (3) speaker score for speaker track annotation (SF), and (4) speaker score for speaker annotation (SS). An active learning cycle, which is depicted as $step$, asks one annotation for each video. As the performance measure, we use the F-measure, instead of accuracy, since the number of instances of each class (person) in the datasets is very different.

We have designed two different evaluation techniques to measure the performance of the algorithms. The first evaluation technique calculates overall performance for all videos by admitting 28 videos as one big video combining speaker and head tracks from all videos. We have then calculated FF, FS, SF, and SS scores (Figure 5.1, 5.2, 5.3 and 5.4).

The multimodal classifying problems FS and SF as seen in Figure 5.2 and 5.3 have more deviations than the others because the correlation between face tracks to speaker tracks is weaker than head to head tracks or speaker to speaker tracks. Therefore, in FS, MDS-Certainty performs better than all others, and more robust than BCS.

In the FS experiment, BCS selects good annotations at the earlier steps, and achieves high F-measure value faster than MDS-Certainty since there is no sufficient data early on to be disagreed on the MDS-Certainty method. However, at the later steps, F-measure value decreases, but MDS-Certainty keeps increasing. To take advantage of

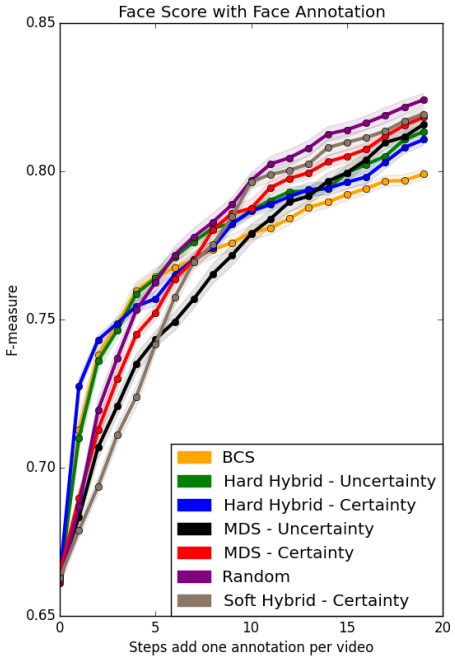


Figure 5.1: Face score with face annotation.

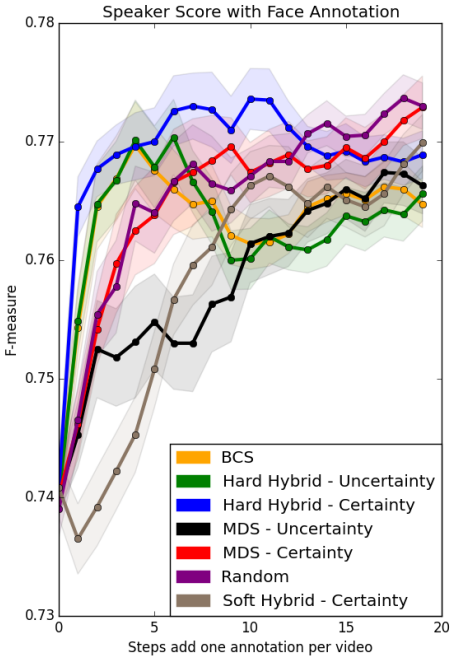


Figure 5.2: Speaker score with face annotation.

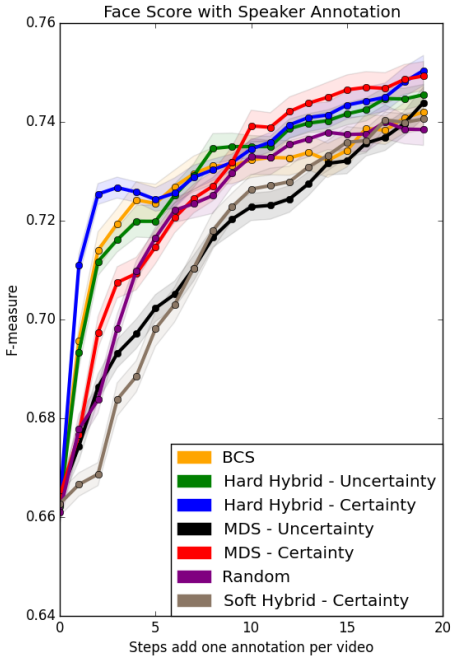


Figure 5.3: Face score with with speaker annotation.

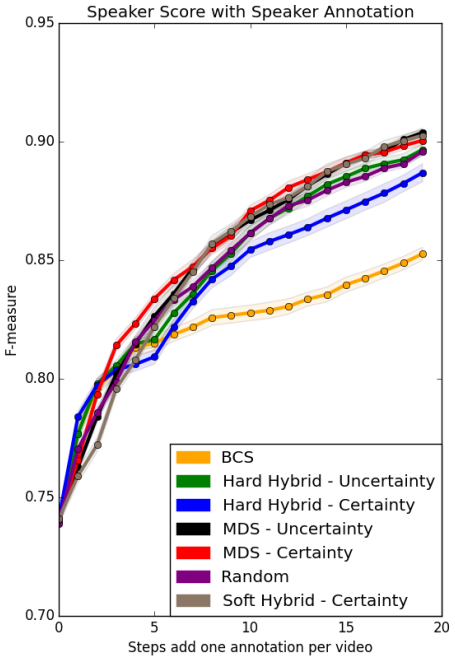


Figure 5.4: Speaker score with with speaker annotation.

the strongest sides of both methods, Hard Hybrid Certainty (HHC) uses BCS at the first five steps. After the fifth step, it uses MDS and its F-measure score keeps increasing. As a result, HHC achieves better scores than random.

In the SF experiment, BCS achieves high F-measure scores rapidly and higher than random method like in FS. Nevertheless, it decreases at the later steps, and finally performs worse than random method. On the other hand, MDS-Certainty F-measure increases stably like that of FS, but it performs similar to random. However, the most interesting result comes from HHC that combines BCS and MDS-Certainty. Like BCS, HHC gives higher F-measures rapidly at the earlier steps; however, it continues to increase at the later stages like MDS-Certainty. Furthermore, it performs significantly better than random until the 10th step. Fortunately, for active learning, reaching to a high F-measure score rapidly is more important.

The second evaluation calculates F-measure scores for each video individually and returns one F-measure value that is the average of all F-measures. By doing so, we observe the performance of the algorithms video by video that have different size of similarity matrices. The video based F-measure results are given in **??**for the four types of experiments.

We have noticed that some MDS based methods perform well when we use speaker annotation; especially HHC performs better for speaker identification with speaker annotation. We have calculated average values of video based performances. The results are given in Figure 5.5, 5.6, 5.7 and 5.8.

FF result in Figure 5.5 shows us HHC is barely better than other methods, and worse than BCS for SF experiment in Figure 5.6. On the other hand, HHC scores are better than speaker annotated experiments with FS and SS shown in Figure 5.7 and Figure 5.8. Another finding is that the deviations are high between video performances. For example, F-measure scores for some videos are between 0.7-0.9 while they are zero for some other videos. Such deviations indicate that we lose some information about the performance analysis of the methods using average video results. We have also ranked between active learning methods performance step by step for each videos, and created heat maps for all videos and all experiment types (Appendix A.1). We have calculated the average ranks using video based F-measure score averages, and average rankings of the methods for discussion since we apply active learning methods to different video types.

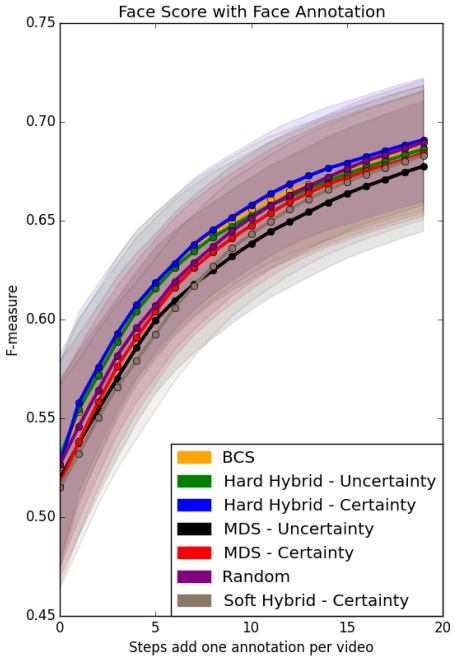


Figure 5.5: Face score with face annotation

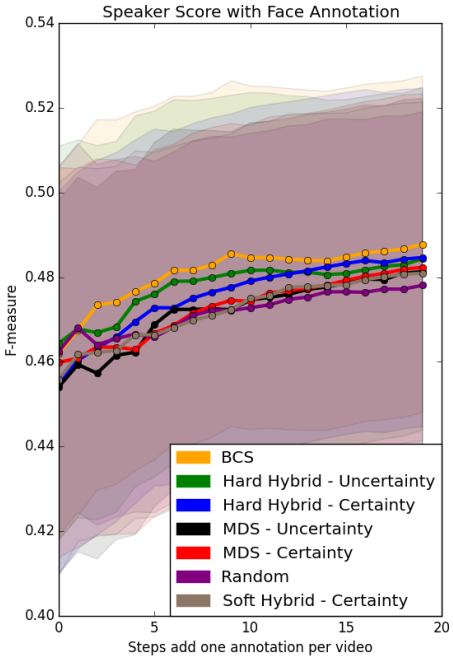


Figure 5.6: Speaker score with face annotation.

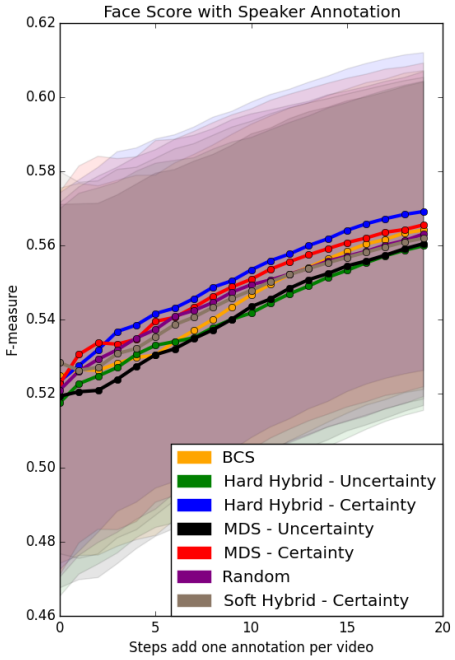


Figure 5.7: Face score with with speaker annotation.

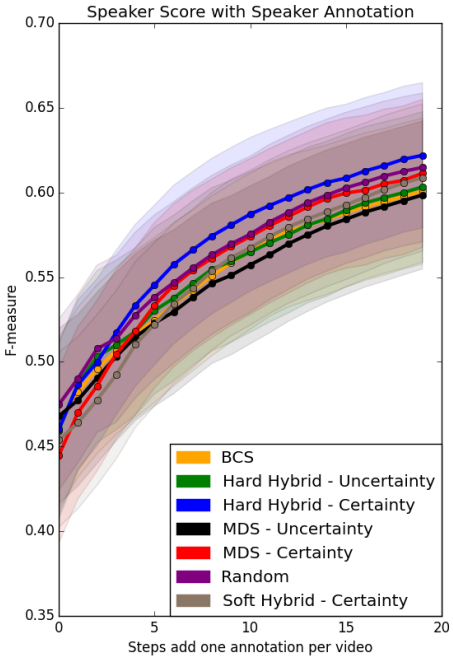


Figure 5.8: Speaker score with with speaker annotation.

Our first finding is that HHC and HHU have the best performance rank average with BCS, but random selection beats BCS at late steps in FF experiment. On the other hand, both MDS-Certainty and MDS-Uncertainty performances are worse than

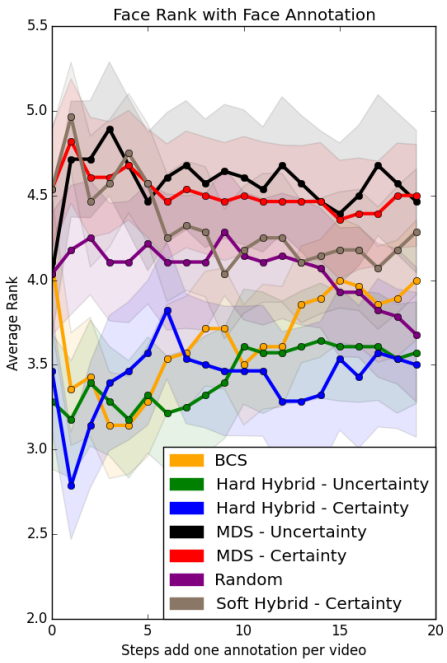


Figure 5.9: Face score with face annotation (lower is better).

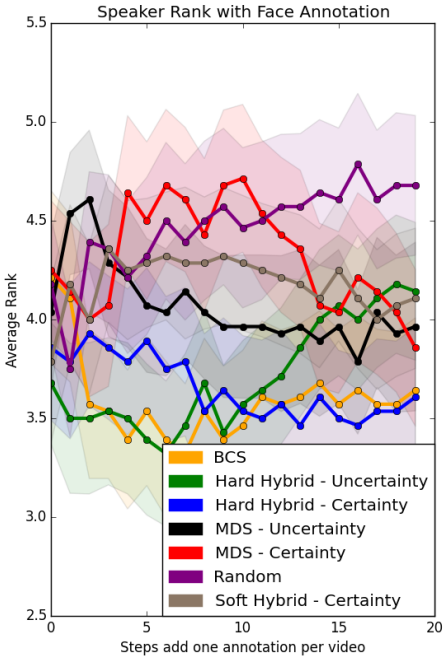


Figure 5.10: Speaker score with face annotation (lower is better).

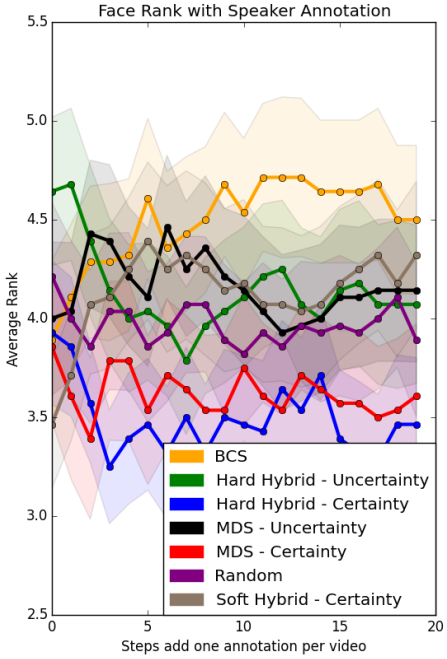


Figure 5.11: Face score with with speaker annotation (lower is better).

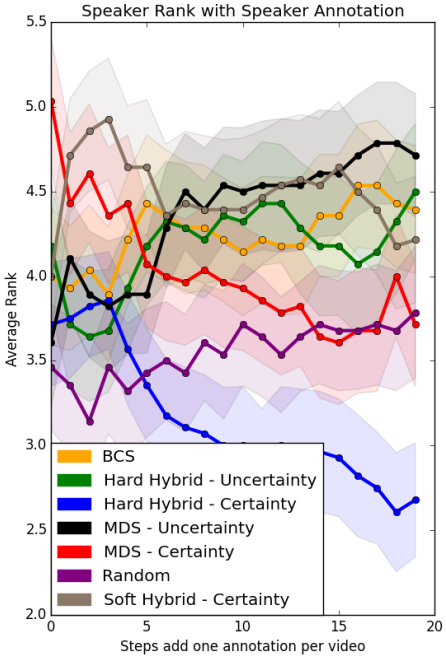


Figure 5.12: Speaker score with with speaker annotation (lower is better).

others. By examining the similarity matrices, we have found out that face-to-face average similarities are approximately ten times lower than speaker-to-speaker average similarities. Both MDS methods do not perform well because of lack of information about the data distribution for disagreement. Hard Hybrid (HH) method performs better due to BCS and MDS since at first 5 steps HH uses BCS, which collects sufficient information on early steps, and MDS which works better after the fifth step.

BCS has better average F-measure score, and ties with HHC in average ranking but all active learning methods beat random selection in SF experiment (Figure 5.10 and 5.10). On the contrary, HHC and MDS-Certainty outperform BCS and random in terms of both average F-measure score and average rankings (Figure 5.11 and 5.11). When we use cross modal annotation, total performance increments are less than single modality.

We know that speaker-to-speaker similarities higher than face-to-face similarities. HHC gives best performance in SS experiment (Figure 5.12) because of higher speaker similarities. We conclude that HHC is better than other sampling methods when similarities between annotated instances are high. High track similarity provides better clustering, and better clustering provides sharp disagreement or agreement. This is why MDS based hybrid method works well.

6. CONCLUSIONS AND RECOMMENDATIONS

Active learning mainly aims to reduce annotation/labeling cost. Active learning is a research area in machine learning and active learning is semi-supervised method. Main idea of the active learning is finding most informative instances for annotating or labeling train data. The most informative instance means that instance gives more information about data distribution so learner can learn data distribution quickly than supervised learning in theory. Annotation cost is important for some cases like video annotation because annotator spends long time and attention to labeling the video so this process is very costly and researchers research active learning algorithms to solving annotation cost problem. In addition, We have big data today especially Internet content like social networks, blogs, video sharing sites etc. annotation cost is high for many cases for big data because this is big, noisy, sparse this is why big data needs more annotation and active learning can reduce cost of big data annotation.

We proposed the MDS active learning method and its hybrid variations and we applied them on multimodal video annotation data. According to our experiments, for different types of annotation tasks, different active learning strategies could be more suitable. Hybrid strategies could be more successful than using a single strategy alone. Decision of the cluster selection and instance selection method adaptively during each active learning step, using a synthetic dataset to investigate the merits of these strategies, examination of each method for each video, rather than the whole REPERE corpus are the future research directions we aim to follow.

REFERENCES

[1] **Giraudel, A.**, **Carré, M.**, **Mapelli, V.**, **Kahn, J.**, **Galibert, O. and Quintard, L.**, 2012. The REPERE Corpus: a multimodal corpus for person recognition., LREC, pp.1102–1107.

[2] **Budnik, M.**, **Poignant, J.**, **Besacier, L.**, **Quénot, G.** *et al.*, 2014. Automatic propagation of manual annotations for multimodal person identification, Proceeding of the 12th International Workshop on Content-Based Multimedia Indexing.

[3] YouTube Statistics 2014, `https://www.youtube.com/yt/press/statistics.html`.

[4] **Ayache, S. and Quénot, G.**, 2007. Evaluation of active learning strategies for video indexing, *Signal Processing: Image Communication*, **22(7)**, 692–704.

[5] **Ayache, S. and Quénot, G.**, 2008. Video corpus annotation using active learning, Advances in Information Retrieval, Springer, pp.187–198.

[6] **Joshi, H.**, **Bayrak, C. and Xu, X.**, 2006. UALR at TREC: Blog Track., TREC.

[7] **Bilgic, M.**, **Mihalkova, L. and Getoor, L.**, 2010. Active learning for networked data, Proceedings of the 27th International Conference on Machine Learning (ICML-10), pp.79–86.

[8] **Raghavan, H.**, **Madani, O. and Jones, R.**, 2006. Active learning with feedback on features and instances, *The Journal of Machine Learning Research*, **7**, 1655–1686.

[9] Collaborative Annotation of multi-Modal, multI-Lingual and multi-media documents, `http://www.chistera.eu/projects/camomile`.

[10] **Poignant, J.**, **Besacier, L.**, **Quénot, G. and Thollard, F.**, 2012. From text detection in videos to person identification, Multimedia and Expo (ICME), 2012 IEEE International Conference on, IEEE, pp.854–859.

[11] **Anthimopoulos, M.**, **Gatos, B. and Pratikakis, I.**, 2010. A two-stage scheme for text detection in video images, *Image and Vision Computing*, **28(9)**, 1413–1426.

[12] **Chen, S. and Gopalakrishnan, P.**, 1998. Speaker, environment and channel change detection and clustering via the Bayesian Information Criterion, Proc. DARPA Broadcast News Transcription and Understanding Workshop, Virginia, USA, p. 8.

[13] **Guillaumin, M.**, **Mensink, T.**, **Verbeek, J. and Schmid, C.**, 2012. Face recognition from caption-based supervision, *International Journal of Computer Vision*, **96(1)**, 64–82.

[14] **Settles, B.**, 2012. Active learning, *Synthesis Lectures on Artificial Intelligence and Machine Learning*, **6(1)**, 1–114.

[15] **Valiant, L.G.**, 1984. A theory of the learnable, *Communications of the ACM*, **27(11)**, 1134–1142.

[16] **Angluin, D.**, 1988. Queries and concept learning, *Machine learning*, **2(4)**, 319–342.

[17] **Settles, B.**, 2010. Active learning literature survey, *University of Wisconsin, Madison*, **52**, 55–66.

[18] **Lewis, D.D. and Catlett, J.**, 1994. Heterogenous Uncertainty Sampling for Supervised Learning., ICML, volume 94, pp.148–156.

[19] **Shannon, C.E.**, 2001. A mathematical theory of communication, *ACM SIGMOBILE Mobile Computing and Communications Review*, **5(1)**, 3–55.

[20] **Cohn, D.**, **Atlas, L. and Ladner, R.**, 1994. Improving generalization with active learning, *Machine learning*, **15(2)**, 201–221.

[21] **Seung, H.S.**, **Opper, M. and Sompolinsky, H.**, 1992. Query by committee, Proceedings of the fifth annual workshop on Computational learning theory, ACM, pp.287–294.

[22] **Breiman, L.**, 1996. Bagging predictors, *Machine learning*, **24(2)**, 123–140.

[23] **Kullback, S. and Leibler, R.A.**, 1951. On information and sufficiency, *The Annals of Mathematical Statistics*, 79–86.

[24] **Meilă, M.**, 2003. Comparing clusterings by the variation of information, Learning theory and kernel machines, Springer, pp.173–187.

[25] **Wallace, D.L.**, 1983. Comment, *Journal of the American Statistical Association*, **78(383)**, 569–576.

[26] **Rand, W.M.**, 1971. Objective criteria for the evaluation of clustering methods, *Journal of the American Statistical association*, **66(336)**, 846–850.

[27] **Mirkin, B.**, 1998. Mathematical classification and clustering: From how to what and why, Springer.

[28] **Meilă, M. and Heckerman, D.**, 2001. An experimental comparison of model-based clustering methods, *Machine learning*, **42(1-2)**, 9–29.

[29] **Dongen, S.**, 2000. Performance criteria for graph clustering and Markov cluster experiments.

[30] **Amigó, E.**, **Gonzalo, J.**, **Artiles, J. and Verdejo, F.**, 2009. A comparison of extrinsic clustering evaluation metrics based on formal constraints, *Information retrieval*, **12(4)**, 461–486.

[31] **Iwama, K. and Miyazaki, S.**, 2008. A survey of the stable marriage problem and its variants, Informatics Education and Research for Knowledge-Circulating Society, 2008. ICKS 2008. International Conference on, IEEE, pp.131–136.

[32] **Krumpelman, C. and Ghosh, J.**, 2007. Matching and visualization of multiple overlapping clusterings of microarray data, Computational Intelligence and Bioinformatics and Computational Biology, 2007. CIBCB'07. IEEE Symposium on, IET, pp.121–126.

[33] **Cohn, D.A.**, **Ghahramani, Z. and Jordan, M.I.**, 1996. Active learning with statistical models, *arXiv preprint cs/9603104*.

[34] **Atlas, L.E.**, **Cohn, D.A. and Ladner, R.E.**, 1990. Training connectionist networks with queries and selective sampling, Advances in neural information processing systems, pp.566–573.

[35] **Settles, B. and Craven, M.**, 2008. An analysis of active learning strategies for sequence labeling tasks, Proceedings of the Conference on Empirical Methods in Natural Language Processing, Association for Computational Linguistics, pp.1070–1079.

[36] **Xu, Z.**, **Akella, R. and Zhang, Y.**, 2007. Incorporating diversity and density in active learning for relevance feedback, Springer.

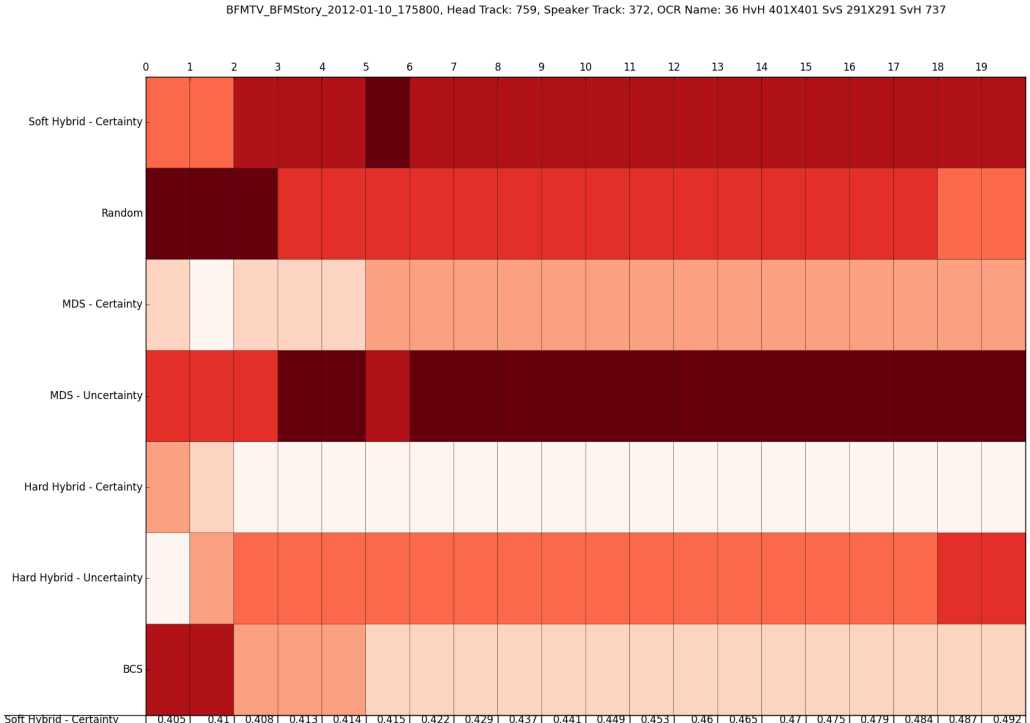
APPENDICES

APPENDIX A.1 : Video Based Results and Methods Ranking Heat Maps

APPENDIX A.1

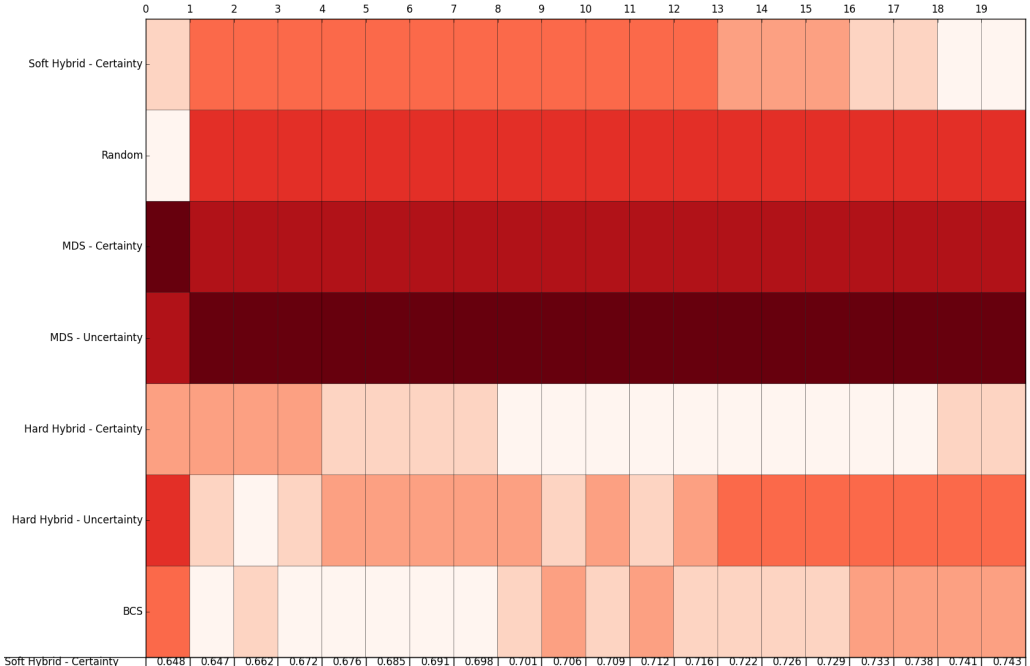
F-measure values and algorithm heat maps of videos.

Face score using face annotation is given in red heat maps.



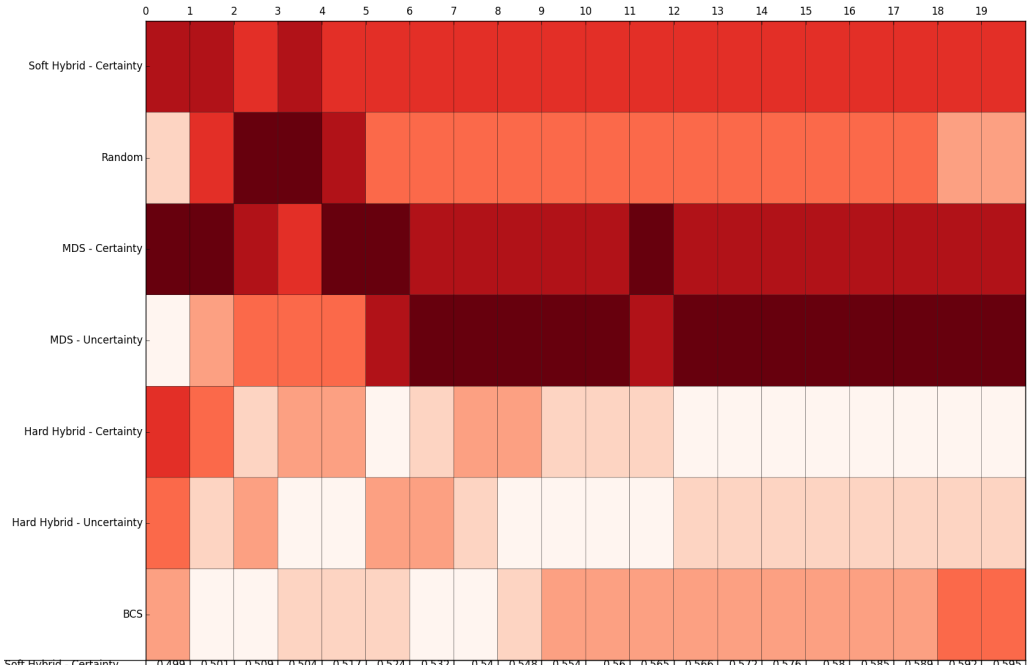
Soft Hybrid - Certainty	0.405	0.41	0.408	0.413	0.414	0.415	0.422	0.429	0.437	0.441	0.449	0.453	0.46	0.465	0.47	0.475	0.479	0.484	0.487	0.492
Random	0.352	0.386	0.404	0.417	0.432	0.436	0.448	0.452	0.462	0.466	0.47	0.474	0.478	0.481	0.488	0.493	0.497	0.5	0.506	0.509
MDS - Certainty	0.421	0.433	0.439	0.441	0.451	0.454	0.456	0.466	0.472	0.478	0.484	0.491	0.494	0.496	0.497	0.5	0.502	0.506	0.51	0.514
MDS - Uncertainty	0.403	0.406	0.408	0.408	0.407	0.415	0.417	0.419	0.42	0.424	0.433	0.436	0.441	0.444	0.449	0.458	0.462	0.468	0.472	0.476
Hard Hybrid - Certainty	0.406	0.43	0.447	0.455	0.465	0.471	0.478	0.489	0.497	0.5	0.503	0.508	0.511	0.517	0.519	0.522	0.526	0.53	0.531	0.535
Hard Hybrid - Uncertainty	0.425	0.415	0.426	0.431	0.436	0.444	0.452	0.46	0.467	0.474	0.483	0.489	0.493	0.492	0.495	0.493	0.498	0.5	0.501	0.503
BCS	0.37	0.404	0.428	0.433	0.45	0.46	0.468	0.473	0.478	0.484	0.49	0.494	0.497	0.5	0.502	0.506	0.51	0.512	0.515	0.518

BFMTV_BFMStory_2012-01-23_175800, Head Track: 782, Speaker Track: 338, OCR Name: 34 HvH 406X406 SvS 247X247 SvH 556



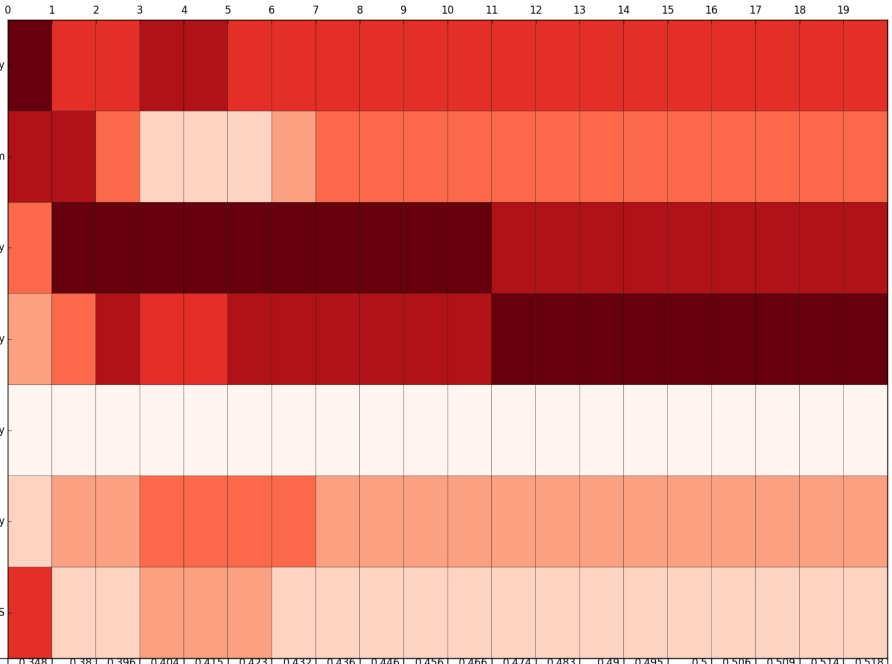
Soft Hybrid - Certainty	0.648	0.647	0.662	0.672	0.676	0.685	0.691	0.698	0.701	0.706	0.709	0.712	0.716	0.722	0.726	0.729	0.733	0.738	0.741	0.743
Random	0.659	0.641	0.638	0.642	0.658	0.666	0.673	0.677	0.684	0.693	0.695	0.702	0.709	0.712	0.715	0.719	0.721	0.723	0.728	0.733
MDS - Certainty	0.581	0.624	0.634	0.641	0.645	0.657	0.664	0.669	0.677	0.681	0.683	0.685	0.69	0.693	0.695	0.699	0.699	0.701	0.705	0.706
MDS - Uncertainty	0.614	0.616	0.625	0.625	0.633	0.648	0.658	0.662	0.665	0.67	0.672	0.677	0.679	0.681	0.686	0.69	0.692	0.695	0.696	0.699
Hard Hybrid - Certainty	0.636	0.654	0.667	0.682	0.7	0.711	0.71	0.718	0.722	0.722	0.726	0.73	0.733	0.735	0.736	0.738	0.739	0.739	0.74	0.742
Hard Hybrid - Uncertainty	0.631	0.663	0.684	0.687	0.692	0.702	0.708	0.714	0.718	0.72	0.722	0.724	0.721	0.722	0.724	0.727	0.726	0.729	0.73	0.734
BCS	0.632	0.674	0.684	0.689	0.702	0.712	0.716	0.719	0.719	0.719	0.723	0.724	0.728	0.728	0.73	0.733	0.733	0.736	0.738	0.737

BFMTV_BFMStory_2012-02-14_175800, Head Track: 753, Speaker Track: 348, OCR Name: 33 HvH 405X405 SvS 272X272 SvH 643



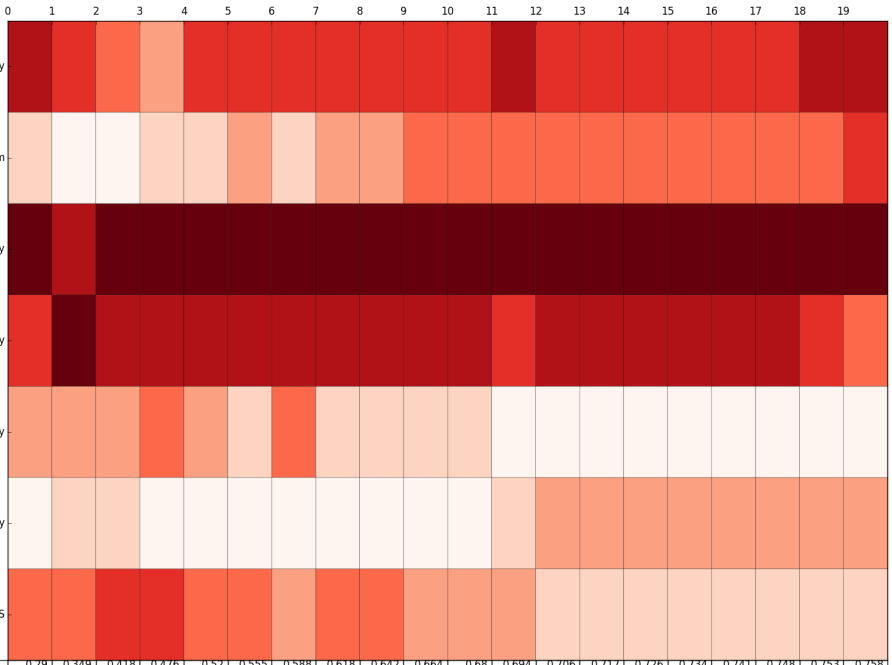
Soft Hybrid - Certainty	0.499	0.501	0.509	0.504	0.517	0.524	0.532	0.54	0.548	0.554	0.56	0.565	0.566	0.572	0.576	0.58	0.585	0.589	0.592	0.595
Random	0.511	0.502	0.495	0.504	0.513	0.526	0.538	0.549	0.553	0.562	0.567	0.574	0.582	0.587	0.591	0.595	0.601	0.605	0.609	0.611
MDS - Certainty	0.487	0.494	0.497	0.506	0.509	0.514	0.524	0.529	0.531	0.535	0.54	0.544	0.553	0.561	0.567	0.57	0.574	0.577	0.581	0.584
MDS - Uncertainty	0.526	0.526	0.523	0.525	0.521	0.523	0.522	0.521	0.525	0.532	0.538	0.546	0.548	0.556	0.562	0.568	0.57	0.575	0.58	0.583
Hard Hybrid - Certainty	0.507	0.526	0.535	0.55	0.555	0.567	0.575	0.578	0.585	0.589	0.595	0.597	0.603	0.606	0.609	0.613	0.615	0.617	0.621	0.624
Hard Hybrid - Uncertainty	0.508	0.531	0.533	0.559	0.567	0.565	0.573	0.58	0.587	0.593	0.598	0.602	0.603	0.603	0.603	0.607	0.61	0.613	0.614	0.618
BCS	0.51	0.533	0.543	0.553	0.564	0.567	0.577	0.582	0.586	0.588	0.59	0.593	0.596	0.597	0.6	0.603	0.605	0.606	0.606	0.607

BFMTV_BFMStory_2012-02-20_175800, Head Track: 871, Speaker Track: 342, OCR Name: 30 HvH 450X450 SvS 267X267 SvH 710



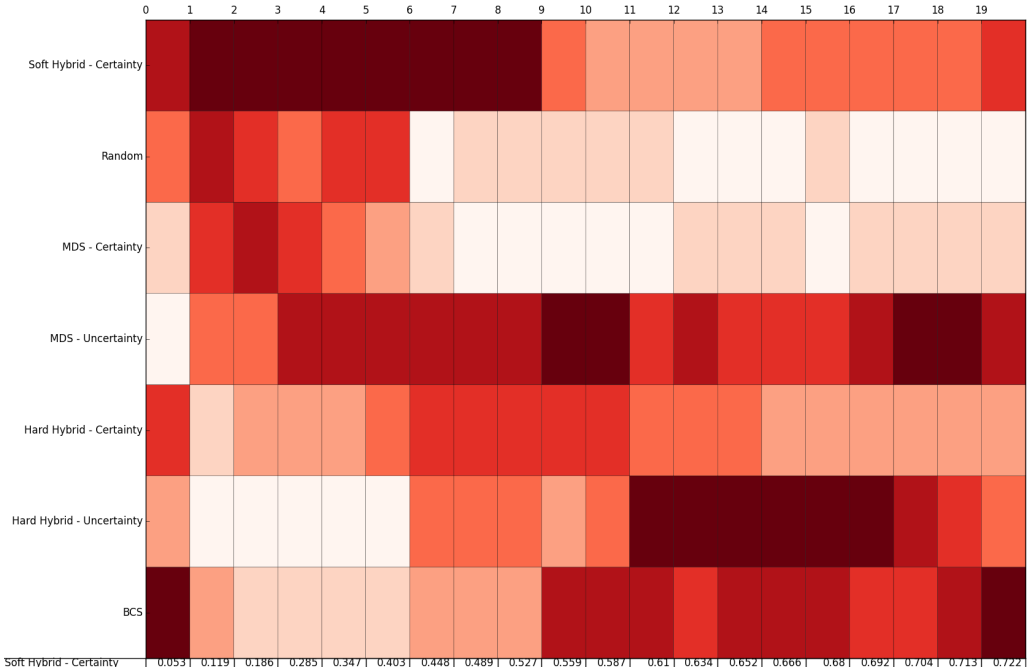
Soft Hybrid - Certainty	0.348	0.38	0.396	0.404	0.415	0.423	0.432	0.436	0.446	0.456	0.466	0.474	0.483	0.49	0.495	0.5	0.506	0.509	0.514	0.518
Random	0.368	0.377	0.399	0.424	0.439	0.453	0.459	0.466	0.473	0.478	0.483	0.488	0.493	0.499	0.504	0.51	0.514	0.517	0.521	0.525
MDS - Certainty	0.375	0.377	0.385	0.387	0.399	0.411	0.422	0.429	0.433	0.439	0.446	0.458	0.469	0.477	0.484	0.49	0.498	0.504	0.511	0.518
MDS - Uncertainty	0.393	0.389	0.394	0.406	0.416	0.418	0.425	0.431	0.439	0.444	0.447	0.452	0.457	0.462	0.467	0.472	0.477	0.483	0.487	0.492
Hard Hybrid - Certainty	0.419	0.433	0.45	0.47	0.482	0.488	0.491	0.495	0.5	0.505	0.512	0.519	0.526	0.529	0.533	0.533	0.537	0.539	0.542	0.546
Hard Hybrid - Uncertainty	0.394	0.391	0.4	0.421	0.432	0.443	0.457	0.471	0.482	0.487	0.489	0.494	0.502	0.506	0.508	0.514	0.519	0.524	0.53	0.534
BCS	0.37	0.392	0.411	0.422	0.438	0.45	0.467	0.478	0.485	0.494	0.501	0.505	0.51	0.516	0.521	0.523	0.527	0.531	0.536	0.539

BFMTV_CultureEtVous_2012-01-13_065040, Head Track: 69, Speaker Track: 38, OCR Name: 3 HvH 29X29 SvS 28X28 SvH 42



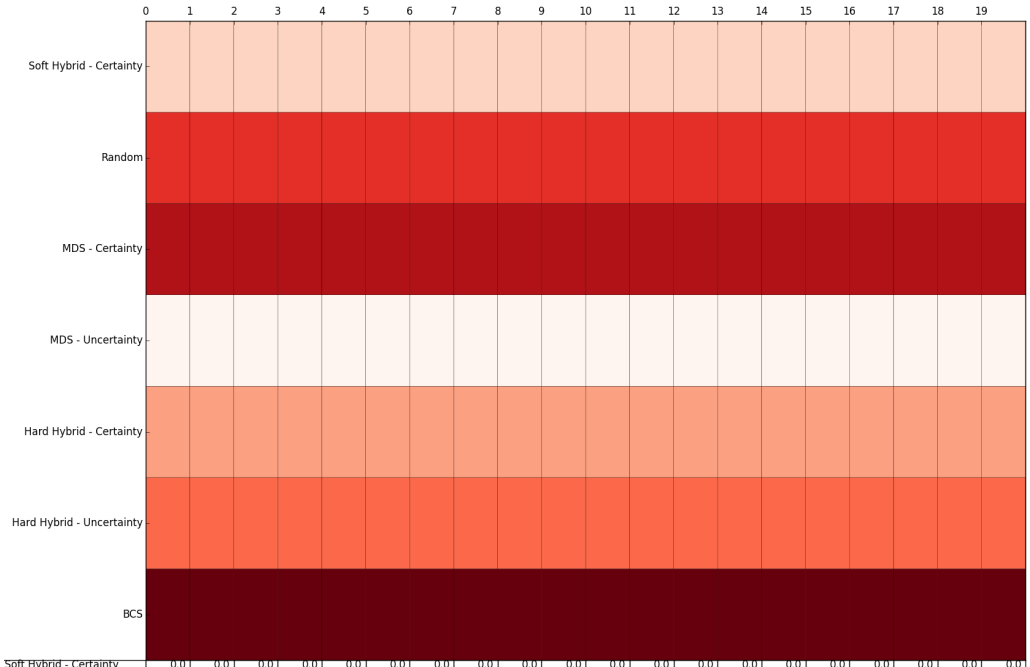
Soft Hybrid - Certainty	0.29	0.349	0.418	0.476	0.52	0.555	0.588	0.618	0.642	0.664	0.68	0.694	0.706	0.717	0.726	0.734	0.741	0.748	0.753	0.758
Random	0.324	0.392	0.449	0.483	0.53	0.572	0.606	0.635	0.658	0.677	0.692	0.705	0.715	0.725	0.733	0.74	0.744	0.75	0.755	0.759
MDS - Certainty	0.289	0.332	0.367	0.426	0.482	0.531	0.567	0.603	0.629	0.652	0.67	0.685	0.698	0.71	0.72	0.728	0.736	0.743	0.749	0.755
MDS - Uncertainty	0.293	0.322	0.372	0.439	0.495	0.54	0.578	0.612	0.638	0.661	0.678	0.694	0.705	0.715	0.724	0.733	0.74	0.747	0.753	0.759
Hard Hybrid - Certainty	0.322	0.36	0.419	0.472	0.528	0.572	0.605	0.635	0.66	0.679	0.695	0.709	0.719	0.729	0.737	0.745	0.751	0.757	0.762	0.767
Hard Hybrid - Uncertainty	0.328	0.372	0.436	0.49	0.544	0.584	0.617	0.643	0.665	0.682	0.697	0.708	0.718	0.727	0.734	0.741	0.747	0.752	0.756	0.76
BCS	0.318	0.357	0.417	0.471	0.527	0.571	0.606	0.634	0.658	0.678	0.694	0.708	0.719	0.729	0.737	0.745	0.751	0.757	0.762	0.767

BFMTV_CultureEtVous_2012-01-16_075000, Head Track: 163, Speaker Track: 42, OCR Name: 5 HvH 78X78 SvS 32X32 SvH 117



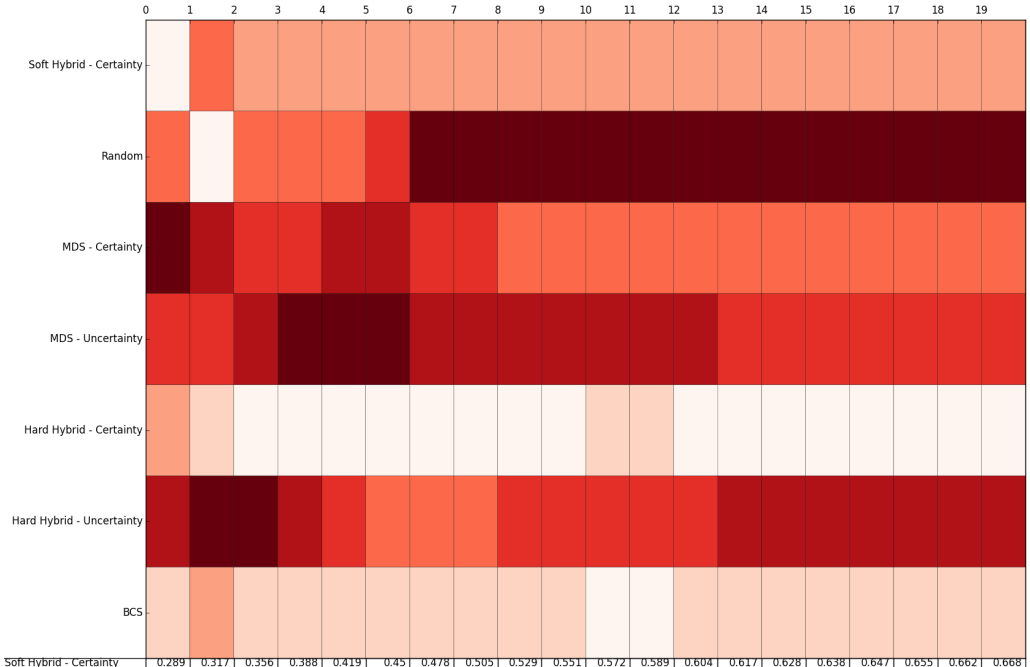
Soft Hybrid - Certainty	0.053	0.119	0.186	0.285	0.347	0.403	0.448	0.489	0.527	0.559	0.587	0.61	0.634	0.652	0.666	0.68	0.692	0.704	0.713	0.722
Random	0.07	0.154	0.261	0.35	0.396	0.445	0.499	0.528	0.556	0.583	0.611	0.632	0.653	0.667	0.684	0.694	0.708	0.718	0.726	0.736
MDS - Certainty	0.092	0.207	0.259	0.34	0.399	0.449	0.496	0.54	0.566	0.596	0.614	0.637	0.653	0.666	0.683	0.695	0.708	0.717	0.724	0.734
MDS - Uncertainty	0.106	0.21	0.284	0.335	0.383	0.417	0.469	0.501	0.528	0.552	0.581	0.607	0.624	0.644	0.662	0.679	0.69	0.698	0.71	0.719
Hard Hybrid - Certainty	0.065	0.255	0.322	0.372	0.412	0.446	0.473	0.502	0.53	0.558	0.585	0.609	0.631	0.65	0.668	0.684	0.698	0.708	0.717	0.726
Hard Hybrid - Uncertainty	0.082	0.283	0.338	0.391	0.425	0.454	0.479	0.506	0.536	0.563	0.586	0.603	0.621	0.641	0.659	0.675	0.69	0.703	0.713	0.723
BCS	0.053	0.254	0.329	0.384	0.422	0.452	0.483	0.512	0.539	0.555	0.582	0.606	0.625	0.644	0.662	0.677	0.691	0.704	0.711	0.716

BFMTV_CultureEtVous_2012-01-17_065040, Head Track: 48, Speaker Track: 49, OCR Name: 5 HvH 23X23 SvS 35X35 SvH 30



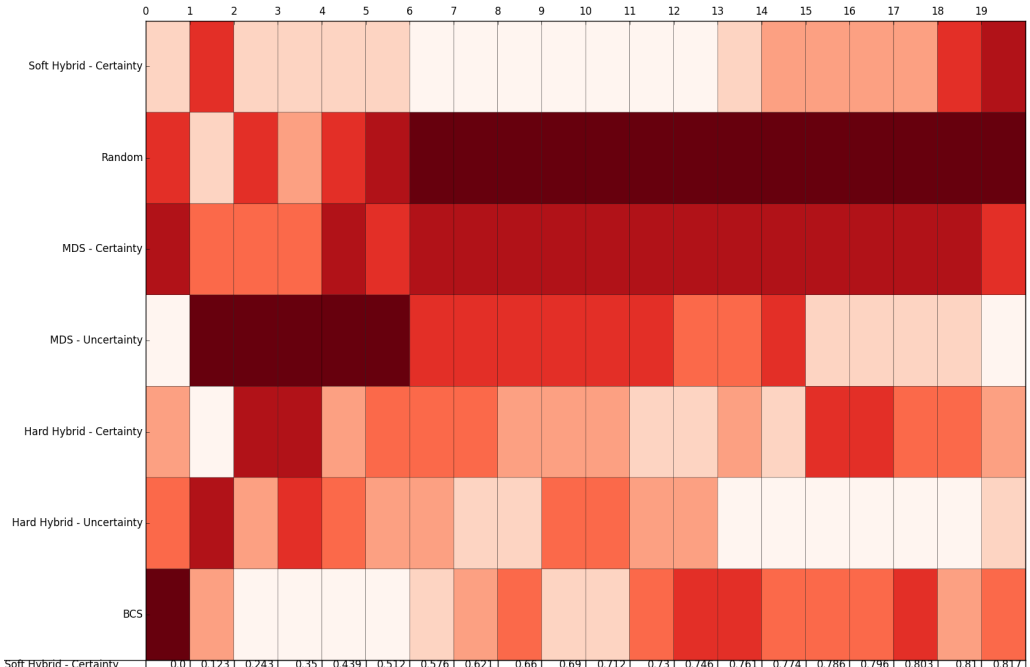
Soft Hybrid - Certainty	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Random	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
MDS - Certainty	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
MDS - Uncertainty	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Hard Hybrid - Certainty	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Hard Hybrid - Uncertainty	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
BCS	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

BFMTV_CultureEtVous_2012-01-18_065040, Head Track: 102, Speaker Track: 47, OCR Name: 6 HvH 41X41 SvS 37X37 SvH 42



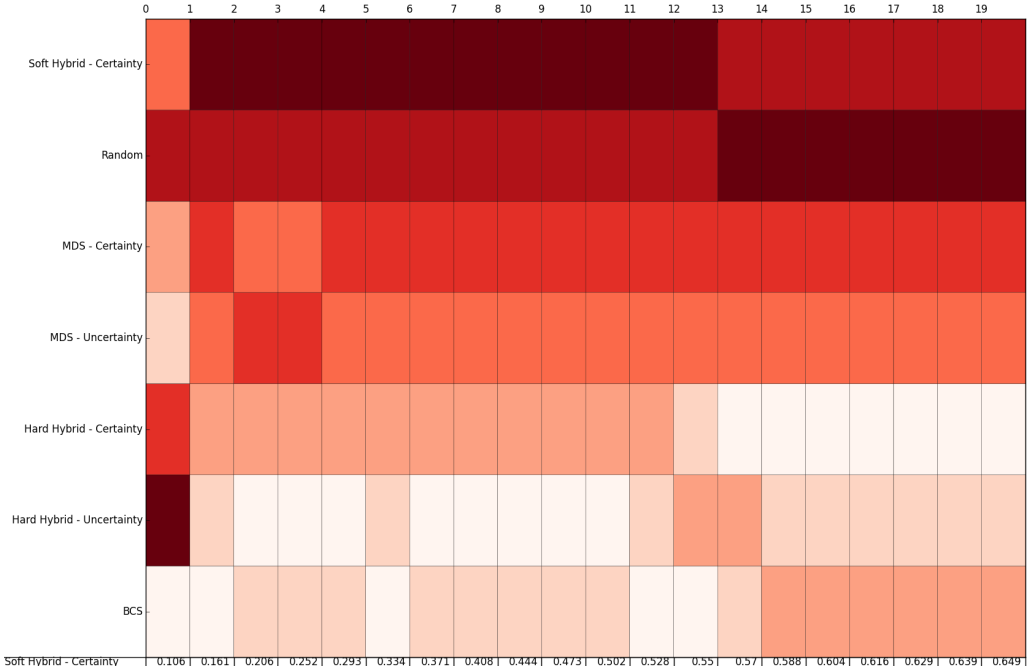
Soft Hybrid - Certainty	0.289	0.317	0.356	0.388	0.419	0.45	0.478	0.505	0.529	0.551	0.572	0.589	0.604	0.617	0.628	0.638	0.647	0.655	0.662	0.668
Random	0.273	0.327	0.351	0.382	0.411	0.439	0.464	0.488	0.507	0.526	0.546	0.563	0.577	0.59	0.601	0.61	0.618	0.626	0.632	0.638
MDS - Certainty	0.232	0.306	0.344	0.378	0.406	0.439	0.469	0.497	0.522	0.546	0.567	0.585	0.6	0.614	0.625	0.636	0.645	0.652	0.659	0.666
MDS - Uncertainty	0.268	0.306	0.338	0.37	0.404	0.434	0.464	0.491	0.516	0.541	0.562	0.58	0.595	0.609	0.621	0.631	0.64	0.649	0.656	0.663
Hard Hybrid - Certainty	0.28	0.325	0.367	0.401	0.43	0.46	0.485	0.511	0.534	0.556	0.575	0.593	0.608	0.621	0.632	0.642	0.65	0.658	0.664	0.67
Hard Hybrid - Uncertainty	0.249	0.288	0.336	0.375	0.407	0.439	0.469	0.497	0.521	0.544	0.564	0.582	0.596	0.609	0.62	0.63	0.639	0.646	0.653	0.659
BCS	0.281	0.321	0.361	0.393	0.424	0.453	0.481	0.508	0.533	0.555	0.576	0.594	0.608	0.621	0.632	0.641	0.65	0.658	0.664	0.67

BFMTV_CultureEtVous_2012-01-19_065040, Head Track: 78, Speaker Track: 41, OCR Name: 4 HvH 43X43 SvS 31X31 SvH 52



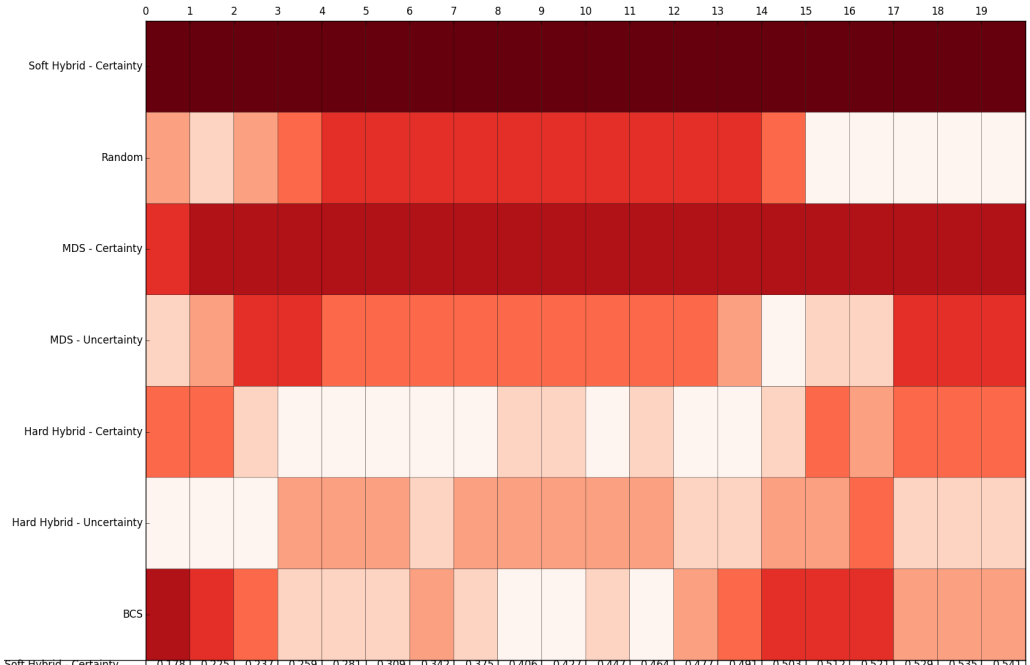
Soft Hybrid - Certainty	0.0	0.123	0.243	0.35	0.439	0.512	0.576	0.621	0.66	0.69	0.712	0.73	0.746	0.761	0.774	0.786	0.796	0.803	0.81	0.817
Random	0.0	0.145	0.234	0.339	0.422	0.487	0.544	0.587	0.621	0.648	0.67	0.689	0.704	0.718	0.729	0.738	0.746	0.754	0.759	0.766
MDS - Certainty	0.0	0.128	0.235	0.336	0.422	0.489	0.548	0.601	0.641	0.674	0.698	0.721	0.741	0.758	0.77	0.782	0.793	0.8	0.81	0.818
MDS - Uncertainty	0.0	0.094	0.205	0.308	0.404	0.484	0.55	0.601	0.643	0.677	0.704	0.726	0.745	0.76	0.773	0.786	0.797	0.807	0.816	0.823
Hard Hybrid - Certainty	0.0	0.149	0.227	0.333	0.424	0.497	0.561	0.609	0.649	0.679	0.705	0.727	0.746	0.761	0.775	0.783	0.793	0.803	0.811	0.819
Hard Hybrid - Uncertainty	0.0	0.122	0.235	0.334	0.424	0.503	0.565	0.613	0.653	0.679	0.705	0.727	0.746	0.762	0.776	0.787	0.798	0.808	0.817	0.822
BCS	0.0	0.141	0.254	0.357	0.447	0.513	0.567	0.612	0.648	0.68	0.706	0.727	0.745	0.76	0.774	0.786	0.796	0.803	0.812	0.819

BFMTV_CultureEtVous_2012-02-14_065040, Head Track: 104, Speaker Track: 48, OCR Name: 9 HvH 48X48 SvS 39X39 SvH 61



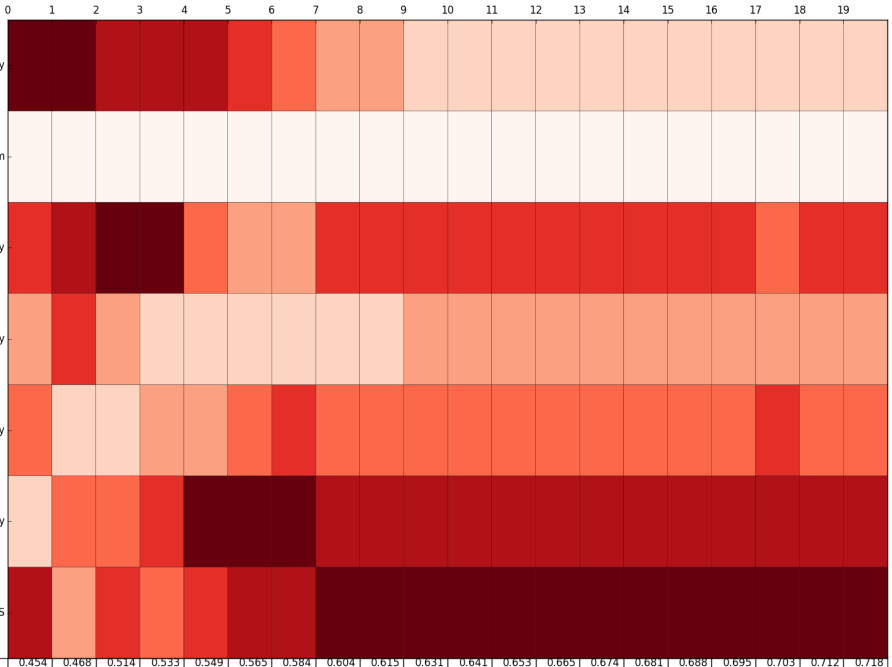
Soft Hybrid - Certainty	0.106	0.161	0.206	0.252	0.293	0.334	0.371	0.408	0.444	0.473	0.502	0.528	0.55	0.57	0.588	0.604	0.616	0.629	0.639	0.649
Random	0.085	0.163	0.225	0.278	0.315	0.357	0.396	0.429	0.461	0.49	0.517	0.537	0.555	0.57	0.585	0.596	0.609	0.618	0.628	0.636
MDS - Certainty	0.11	0.178	0.249	0.306	0.351	0.388	0.424	0.459	0.488	0.516	0.54	0.557	0.575	0.592	0.608	0.62	0.631	0.643	0.65	0.657
MDS - Uncertainty	0.116	0.19	0.243	0.304	0.361	0.401	0.437	0.471	0.497	0.524	0.547	0.567	0.587	0.604	0.617	0.631	0.642	0.65	0.659	0.668
Hard Hybrid - Certainty	0.095	0.221	0.282	0.332	0.375	0.421	0.458	0.486	0.519	0.544	0.564	0.585	0.606	0.622	0.637	0.651	0.663	0.672	0.68	0.688
Hard Hybrid - Uncertainty	0.072	0.222	0.292	0.344	0.389	0.424	0.461	0.494	0.52	0.545	0.567	0.586	0.603	0.62	0.635	0.647	0.657	0.668	0.676	0.685
BCS	0.134	0.232	0.29	0.343	0.383	0.426	0.459	0.49	0.52	0.545	0.567	0.588	0.607	0.622	0.633	0.644	0.655	0.666	0.675	0.683

BFMTV_CultureEtVous_2012-02-15_065040, Head Track: 180, Speaker Track: 34, OCR Name: 4 HvH 67X67 SvS 27X27 SvH 113



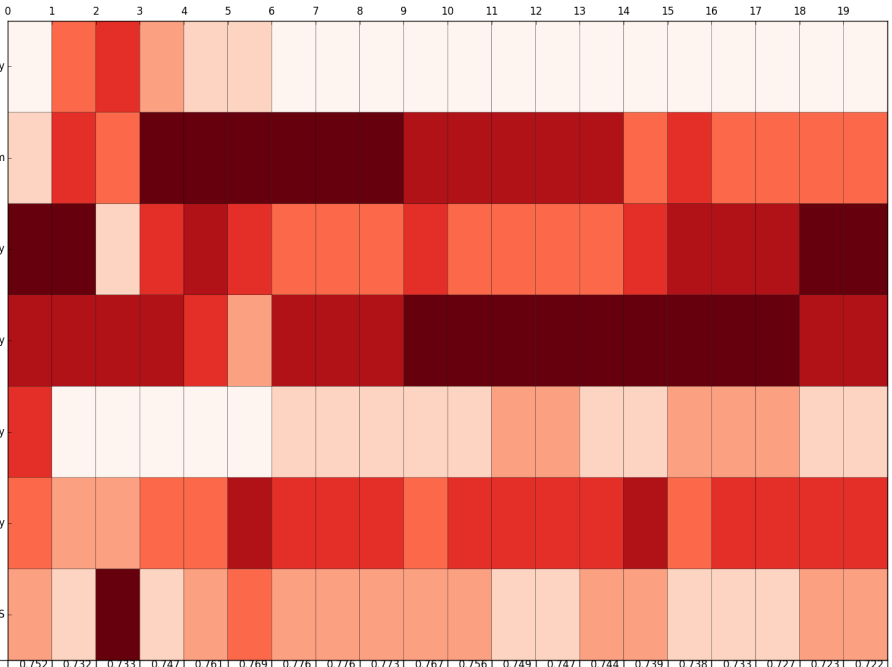
Soft Hybrid - Certainty	0.178	0.225	0.237	0.259	0.281	0.309	0.342	0.375	0.406	0.427	0.447	0.464	0.477	0.491	0.503	0.512	0.521	0.529	0.535	0.541
Random	0.23	0.282	0.33	0.365	0.386	0.408	0.429	0.452	0.468	0.486	0.503	0.519	0.536	0.55	0.562	0.572	0.582	0.59	0.597	0.603
MDS - Certainty	0.218	0.245	0.298	0.333	0.356	0.389	0.419	0.443	0.466	0.483	0.501	0.515	0.528	0.537	0.546	0.554	0.562	0.569	0.575	0.578
MDS - Uncertainty	0.231	0.274	0.316	0.359	0.39	0.418	0.445	0.472	0.491	0.507	0.522	0.535	0.547	0.556	0.565	0.571	0.574	0.575	0.58	0.582
Hard Hybrid - Certainty	0.221	0.271	0.33	0.382	0.423	0.453	0.476	0.496	0.508	0.523	0.536	0.544	0.552	0.56	0.565	0.569	0.574	0.577	0.581	0.583
Hard Hybrid - Uncertainty	0.234	0.286	0.336	0.377	0.413	0.448	0.472	0.491	0.503	0.518	0.531	0.543	0.551	0.558	0.564	0.57	0.574	0.58	0.585	0.589
BCS	0.2	0.259	0.328	0.38	0.421	0.452	0.472	0.492	0.51	0.524	0.535	0.545	0.551	0.556	0.561	0.568	0.573	0.578	0.582	0.586

LCP_CaVousRegarde_2011-12-20_204600, Head Track: 397, Speaker Track: 332, OCR Name: 30 HvH 282X282 SvS 268X268 SvH 406



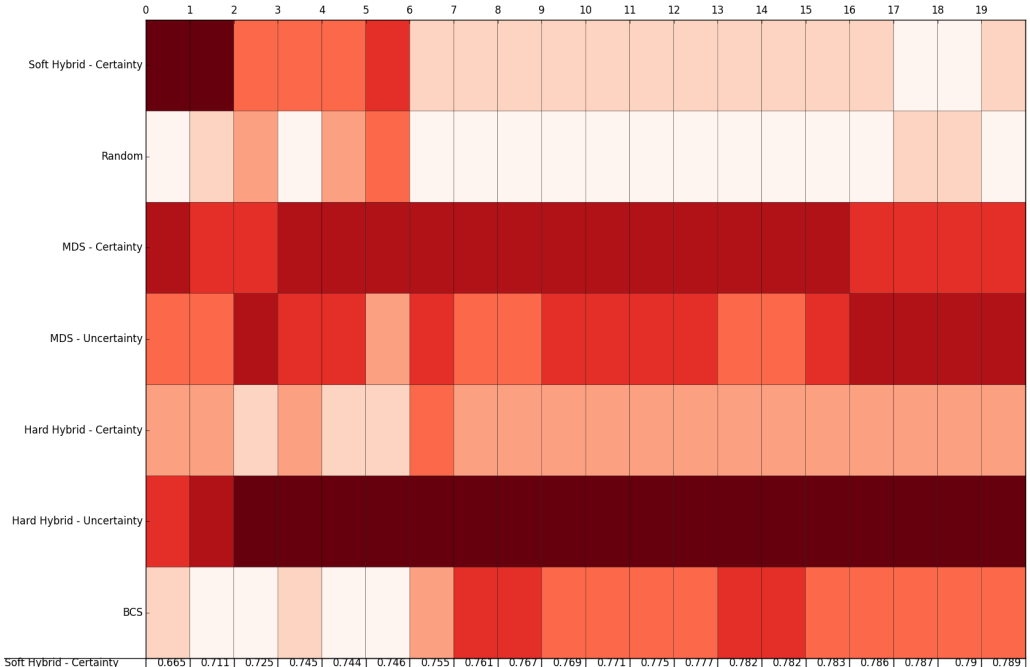
Soft Hybrid - Certainty	0.454	0.468	0.514	0.533	0.549	0.565	0.584	0.604	0.615	0.631	0.641	0.653	0.665	0.674	0.681	0.688	0.695	0.703	0.712	0.718
Random	0.53	0.575	0.589	0.626	0.639	0.644	0.658	0.671	0.679	0.687	0.7	0.712	0.72	0.729	0.736	0.743	0.75	0.757	0.765	0.771
MDS - Certainty	0.463	0.475	0.495	0.528	0.557	0.574	0.592	0.599	0.604	0.611	0.624	0.63	0.64	0.646	0.655	0.66	0.67	0.679	0.686	0.693
MDS - Uncertainty	0.496	0.501	0.546	0.572	0.578	0.593	0.604	0.609	0.617	0.627	0.634	0.645	0.653	0.662	0.669	0.676	0.683	0.69	0.698	0.706
Hard Hybrid - Certainty	0.496	0.52	0.549	0.561	0.571	0.572	0.584	0.601	0.612	0.619	0.627	0.636	0.643	0.652	0.659	0.666	0.671	0.678	0.687	0.694
Hard Hybrid - Uncertainty	0.498	0.502	0.523	0.536	0.547	0.554	0.566	0.583	0.591	0.603	0.612	0.622	0.629	0.642	0.65	0.657	0.666	0.673	0.681	0.687
BCS	0.461	0.503	0.516	0.542	0.557	0.563	0.569	0.574	0.58	0.584	0.589	0.596	0.6	0.605	0.61	0.616	0.626	0.634	0.644	0.654

LCP_CaVousRegarde_2012-01-19_204700, Head Track: 371, Speaker Track: 290, OCR Name: 22 HvH 244X244 SvS 229X229 SvH 384



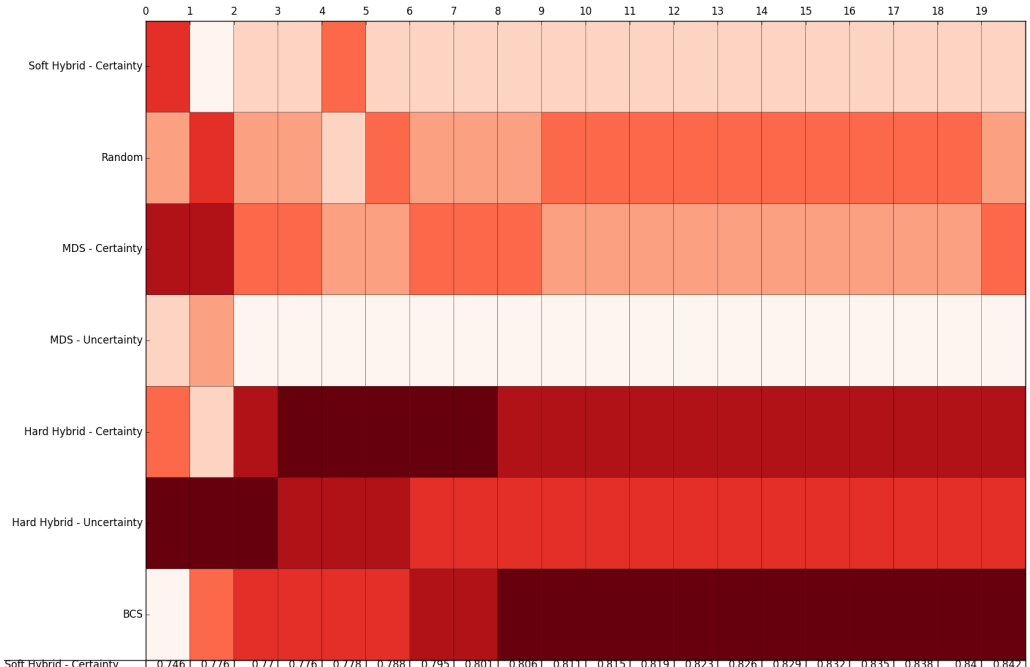
Soft Hybrid - Certainty	0.752	0.732	0.733	0.747	0.761	0.769	0.776	0.776	0.773	0.767	0.756	0.749	0.747	0.744	0.739	0.738	0.733	0.727	0.723	0.722
Random	0.75	0.726	0.735	0.737	0.735	0.729	0.721	0.717	0.717	0.719	0.711	0.703	0.702	0.7	0.701	0.697	0.701	0.699	0.697	0.694
MDS - Certainty	0.663	0.701	0.746	0.739	0.744	0.747	0.745	0.741	0.736	0.723	0.719	0.714	0.71	0.706	0.699	0.692	0.686	0.681	0.675	0.673
MDS - Uncertainty	0.683	0.725	0.731	0.738	0.744	0.75	0.743	0.733	0.724	0.712	0.702	0.702	0.692	0.689	0.681	0.679	0.676	0.675	0.678	0.676
Hard Hybrid - Certainty	0.709	0.744	0.756	0.761	0.77	0.773	0.77	0.761	0.755	0.755	0.747	0.74	0.739	0.731	0.73	0.725	0.721	0.721	0.72	0.719
Hard Hybrid - Uncertainty	0.739	0.742	0.74	0.746	0.746	0.741	0.744	0.739	0.731	0.724	0.717	0.707	0.705	0.702	0.697	0.697	0.694	0.689	0.686	0.687
BCS	0.748	0.743	0.727	0.751	0.753	0.748	0.757	0.756	0.75	0.751	0.747	0.744	0.74	0.731	0.728	0.727	0.723	0.722	0.719	0.716

LCP_CaVousRegarde_2012-01-25_204700, Head Track: 428, Speaker Track: 352, OCR Name: 25 HvH 281X281 SvS 282X282 SvH 429



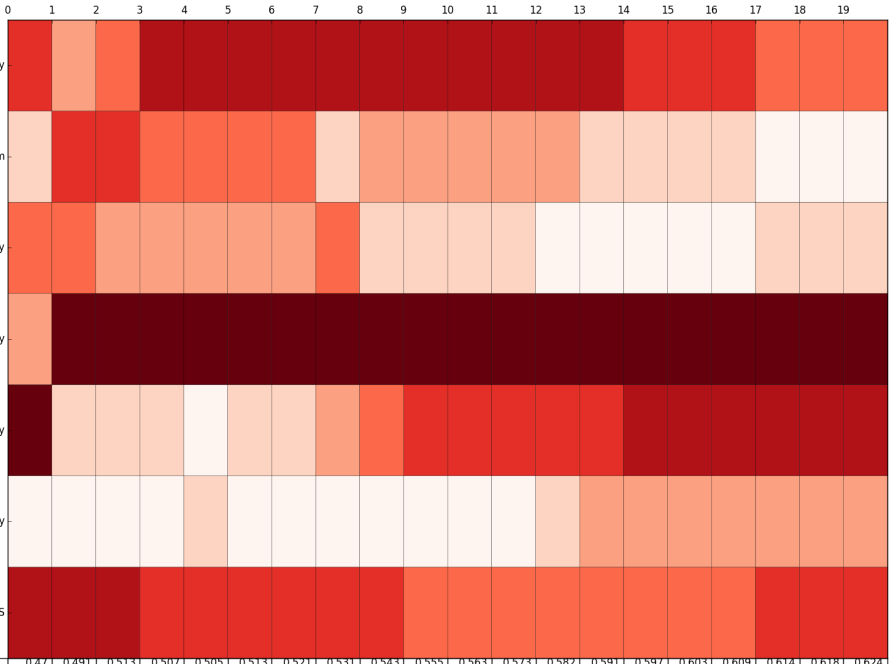
Soft Hybrid - Certainty	0.665	0.711	0.725	0.745	0.744	0.746	0.755	0.761	0.767	0.769	0.771	0.775	0.777	0.782	0.782	0.783	0.786	0.787	0.79	0.789
Random	0.761	0.753	0.732	0.755	0.747	0.747	0.762	0.763	0.769	0.775	0.777	0.778	0.782	0.784	0.783	0.785	0.788	0.787	0.788	0.792
MDS - Certainty	0.676	0.72	0.718	0.728	0.725	0.726	0.733	0.729	0.733	0.738	0.738	0.742	0.749	0.751	0.753	0.755	0.757	0.759	0.762	0.764
MDS - Uncertainty	0.719	0.724	0.716	0.736	0.74	0.751	0.747	0.75	0.749	0.749	0.752	0.752	0.752	0.754	0.757	0.755	0.754	0.756	0.759	0.763
Hard Hybrid - Certainty	0.741	0.736	0.74	0.75	0.756	0.755	0.749	0.753	0.751	0.753	0.755	0.761	0.765	0.764	0.768	0.772	0.772	0.775	0.779	0.782
Hard Hybrid - Uncertainty	0.712	0.715	0.712	0.699	0.708	0.717	0.721	0.711	0.715	0.718	0.717	0.722	0.722	0.727	0.734	0.735	0.741	0.745	0.749	0.752
BCS	0.744	0.755	0.745	0.755	0.758	0.762	0.753	0.749	0.749	0.752	0.755	0.756	0.756	0.754	0.757	0.76	0.765	0.768	0.769	0.769

LCP_EntreLesLignes_2011-12-16_192700, Head Track: 560, Speaker Track: 253, OCR Name: 15 HvH 345X345 SvS 203X203 SvH 493



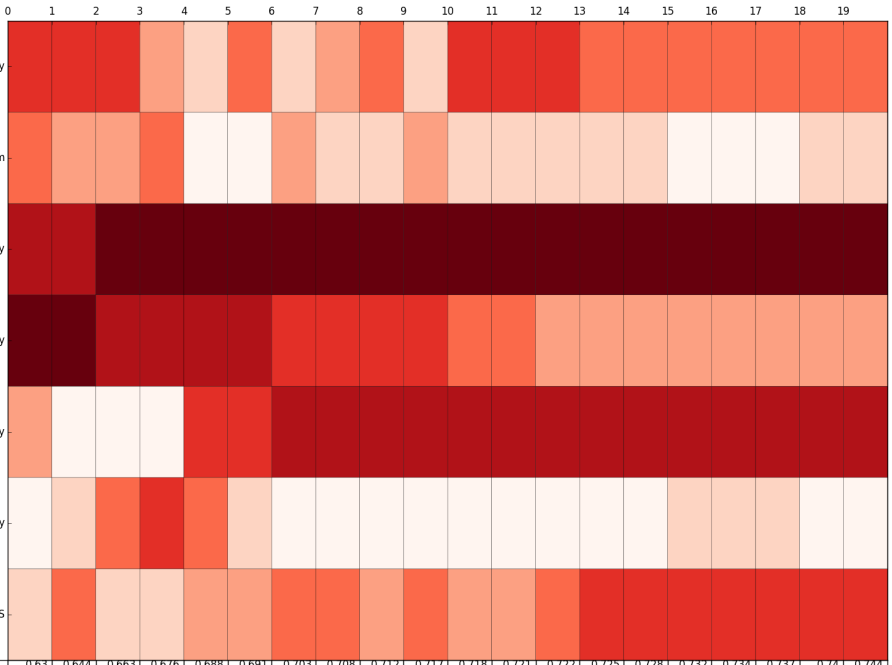
Soft Hybrid - Certainty	0.746	0.776	0.77	0.776	0.778	0.788	0.795	0.801	0.806	0.811	0.815	0.819	0.823	0.826	0.829	0.832	0.835	0.838	0.84	0.842
Random	0.765	0.756	0.768	0.776	0.784	0.786	0.793	0.799	0.801	0.805	0.807	0.812	0.813	0.816	0.821	0.824	0.827	0.831	0.834	0.838
MDS - Certainty	0.719	0.745	0.764	0.772	0.782	0.788	0.792	0.797	0.801	0.807	0.811	0.815	0.82	0.822	0.825	0.828	0.83	0.832	0.835	0.837
MDS - Uncertainty	0.766	0.773	0.776	0.777	0.786	0.792	0.797	0.803	0.808	0.813	0.817	0.821	0.825	0.828	0.831	0.832	0.835	0.838	0.84	0.842
Hard Hybrid - Certainty	0.764	0.774	0.757	0.757	0.751	0.759	0.765	0.773	0.778	0.78	0.784	0.788	0.792	0.796	0.799	0.802	0.805	0.809	0.812	0.814
Hard Hybrid - Uncertainty	0.687	0.731	0.742	0.762	0.762	0.77	0.777	0.785	0.789	0.793	0.8	0.805	0.808	0.813	0.818	0.822	0.826	0.829	0.832	0.835
BCS	0.768	0.763	0.764	0.768	0.771	0.771	0.777	0.779	0.775	0.776	0.776	0.778	0.78	0.784	0.788	0.791	0.794	0.797	0.8	0.803

LCP_EntreLesLignes_2012-01-27_192700, Head Track: 599, Speaker Track: 267, OCR Name: 14 HvH 329X329 SvS 210X210 SvH 460



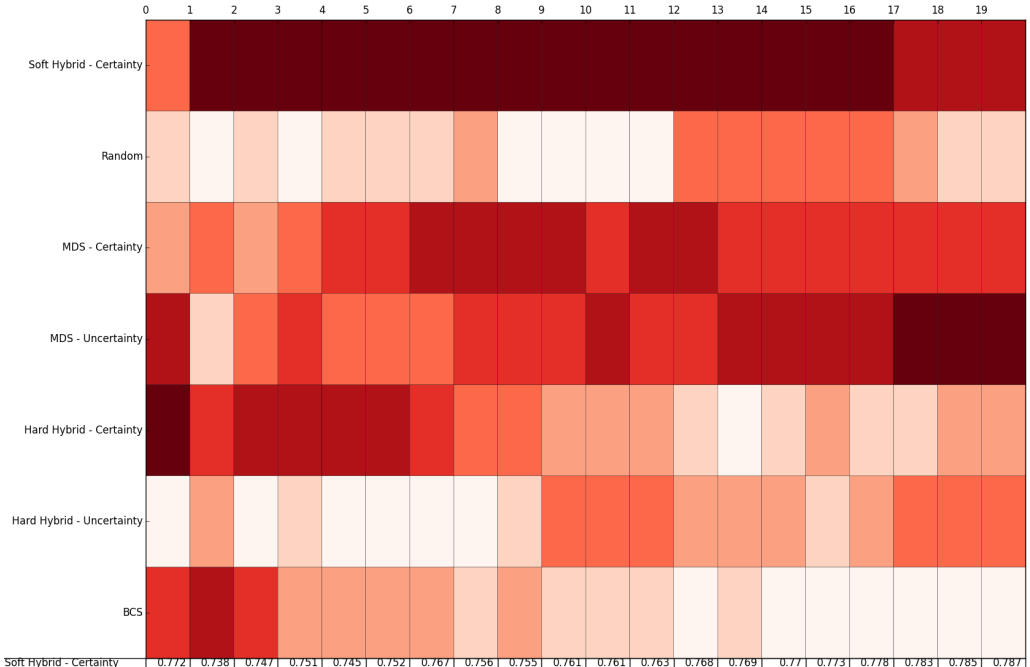
Soft Hybrid - Certainty	0.47	0.491	0.513	0.507	0.505	0.513	0.521	0.531	0.543	0.555	0.563	0.573	0.582	0.591	0.597	0.603	0.609	0.614	0.618	0.624
Random	0.486	0.48	0.504	0.519	0.536	0.543	0.556	0.57	0.573	0.583	0.588	0.596	0.601	0.609	0.615	0.621	0.627	0.633	0.639	0.646
MDS - Certainty	0.473	0.489	0.515	0.528	0.54	0.549	0.56	0.567	0.578	0.584	0.591	0.599	0.605	0.61	0.618	0.623	0.628	0.632	0.636	0.64
MDS - Uncertainty	0.476	0.461	0.47	0.467	0.482	0.498	0.503	0.51	0.514	0.523	0.533	0.539	0.55	0.56	0.569	0.576	0.583	0.591	0.597	0.602
Hard Hybrid - Certainty	0.439	0.503	0.522	0.544	0.554	0.557	0.564	0.569	0.571	0.573	0.577	0.583	0.588	0.592	0.597	0.599	0.603	0.607	0.61	0.613
Hard Hybrid - Uncertainty	0.529	0.526	0.545	0.545	0.553	0.562	0.57	0.577	0.584	0.588	0.593	0.599	0.602	0.607	0.611	0.614	0.62	0.625	0.63	0.635
BCS	0.453	0.477	0.484	0.515	0.535	0.541	0.546	0.557	0.565	0.575	0.582	0.588	0.595	0.599	0.603	0.607	0.611	0.614	0.616	0.619

LCP_EntreLesLignes_2012-05-11_192200, Head Track: 690, Speaker Track: 305, OCR Name: 13 HvH 310X310 SvS 240X240 SvH 596



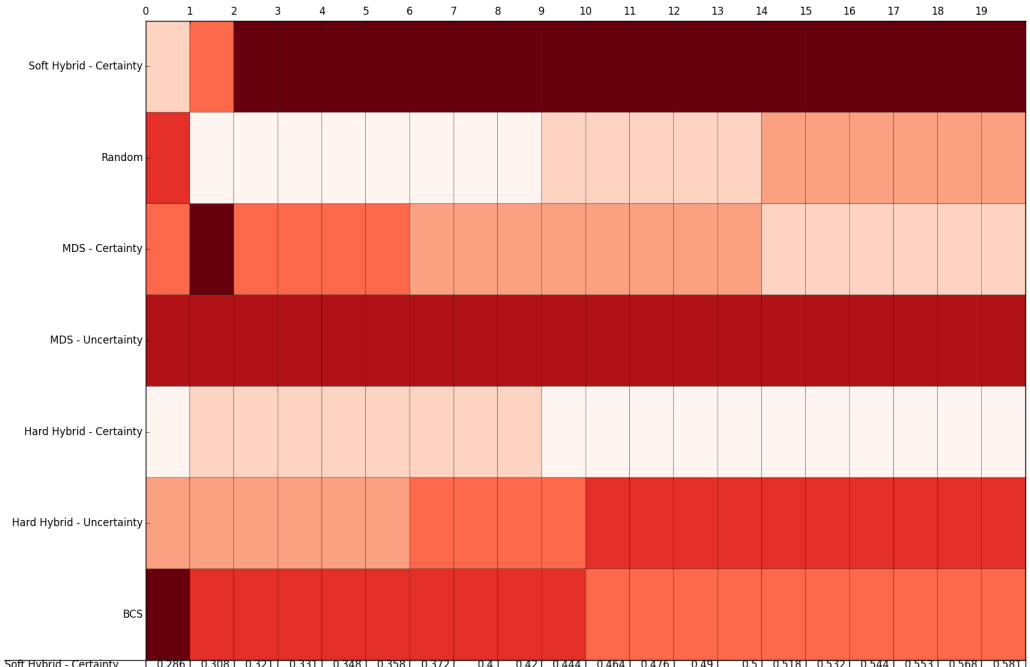
Soft Hybrid - Certainty	0.63	0.644	0.663	0.676	0.688	0.691	0.703	0.708	0.712	0.717	0.718	0.721	0.722	0.725	0.728	0.732	0.734	0.737	0.74	0.744
Random	0.68	0.669	0.671	0.676	0.689	0.7	0.702	0.71	0.715	0.716	0.722	0.728	0.733	0.736	0.74	0.744	0.748	0.749	0.75	0.752
MDS - Certainty	0.611	0.613	0.619	0.621	0.631	0.639	0.651	0.66	0.668	0.672	0.675	0.678	0.682	0.686	0.688	0.692	0.695	0.701	0.705	0.71
MDS - Uncertainty	0.59	0.606	0.627	0.644	0.668	0.683	0.69	0.701	0.708	0.715	0.719	0.721	0.725	0.727	0.731	0.735	0.739	0.742	0.744	0.748
Hard Hybrid - Certainty	0.683	0.685	0.677	0.68	0.683	0.687	0.687	0.691	0.697	0.702	0.704	0.706	0.707	0.708	0.708	0.711	0.715	0.721	0.727	0.731
Hard Hybrid - Uncertainty	0.686	0.683	0.67	0.67	0.684	0.693	0.704	0.713	0.718	0.726	0.729	0.737	0.739	0.74	0.74	0.741	0.742	0.746	0.751	0.752
BCS	0.684	0.658	0.675	0.679	0.685	0.692	0.698	0.708	0.713	0.716	0.722	0.724	0.724	0.723	0.724	0.725	0.727	0.728	0.732	0.734

LCP_LCPInfo13h30_2012-01-24_132700, Head Track: 212, Speaker Track: 163, OCR Name: 13 HvH 128X128 SvS 130X130 SvH 179



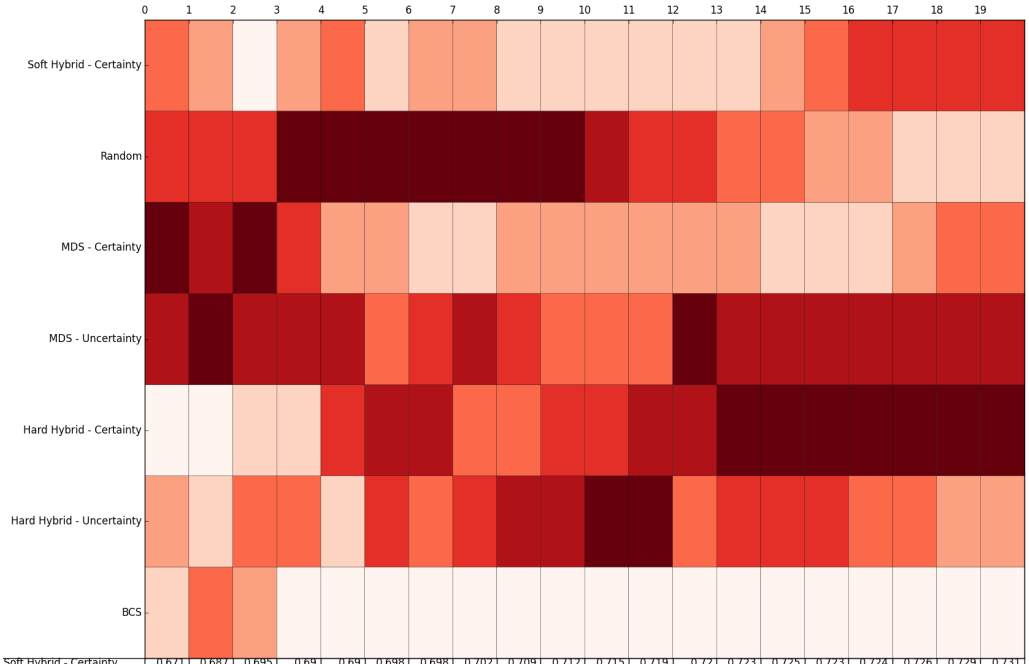
Soft Hybrid - Certainty	0.722	0.738	0.747	0.751	0.745	0.752	0.767	0.756	0.755	0.761	0.761	0.763	0.768	0.769	0.77	0.773	0.778	0.783	0.785	0.787
Random	0.813	0.82	0.802	0.797	0.794	0.796	0.794	0.793	0.803	0.806	0.803	0.806	0.801	0.801	0.799	0.805	0.81	0.814	0.817	0.819
MDS - Certainty	0.794	0.772	0.777	0.772	0.765	0.769	0.77	0.767	0.765	0.769	0.777	0.775	0.774	0.779	0.786	0.788	0.792	0.793	0.797	0.801
MDS - Uncertainty	0.755	0.79	0.767	0.759	0.777	0.781	0.783	0.78	0.776	0.774	0.775	0.78	0.781	0.778	0.781	0.779	0.78	0.779	0.782	0.785
Hard Hybrid - Certainty	0.734	0.764	0.75	0.758	0.757	0.768	0.782	0.789	0.79	0.795	0.797	0.802	0.804	0.809	0.81	0.811	0.812	0.816	0.816	0.818
Hard Hybrid - Uncertainty	0.815	0.79	0.802	0.794	0.803	0.809	0.803	0.801	0.801	0.791	0.794	0.801	0.804	0.808	0.808	0.812	0.812	0.812	0.812	0.815
BCS	0.762	0.756	0.764	0.774	0.78	0.784	0.79	0.796	0.799	0.805	0.803	0.803	0.806	0.809	0.812	0.814	0.816	0.819	0.821	0.822

LCP_LCPInfo13h30_2012-01-25_132700, Head Track: 331, Speaker Track: 172, OCR Name: 10 HvH 208X208 SvS 145X145 SvH 285



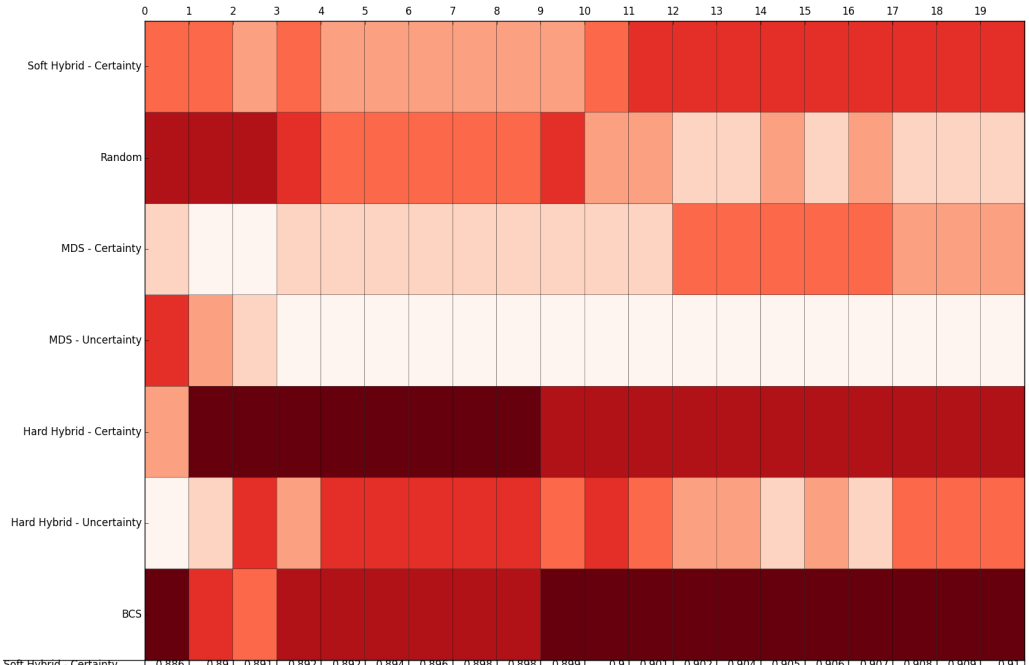
Soft Hybrid - Certainty	0.286	0.308	0.321	0.331	0.348	0.358	0.372	0.4	0.42	0.444	0.464	0.476	0.49	0.5	0.518	0.532	0.544	0.553	0.568	0.581
Random	0.279	0.373	0.419	0.439	0.474	0.494	0.514	0.526	0.541	0.55	0.561	0.574	0.587	0.595	0.602	0.614	0.621	0.63	0.636	0.643
MDS - Certainty	0.281	0.289	0.356	0.39	0.43	0.455	0.481	0.496	0.512	0.526	0.543	0.562	0.574	0.589	0.604	0.615	0.626	0.635	0.646	0.656
MDS - Uncertainty	0.276	0.295	0.336	0.355	0.387	0.412	0.427	0.443	0.449	0.462	0.472	0.481	0.495	0.513	0.527	0.54	0.555	0.565	0.575	0.583
Hard Hybrid - Certainty	0.295	0.332	0.397	0.428	0.456	0.473	0.498	0.519	0.534	0.558	0.577	0.593	0.6	0.612	0.621	0.628	0.638	0.647	0.655	0.662
Hard Hybrid - Uncertainty	0.286	0.318	0.357	0.4	0.436	0.462	0.48	0.489	0.496	0.506	0.52	0.53	0.541	0.552	0.565	0.573	0.579	0.59	0.601	0.609
BCS	0.272	0.306	0.344	0.379	0.42	0.438	0.46	0.474	0.488	0.503	0.521	0.538	0.55	0.566	0.578	0.591	0.603	0.613	0.624	0.634

LCP_LCPInfo13h30_2012-01-27_132700, Head Track: 306, Speaker Track: 177, OCR Name: 11 HvH 182X182 SvS 135X135 SvH 270



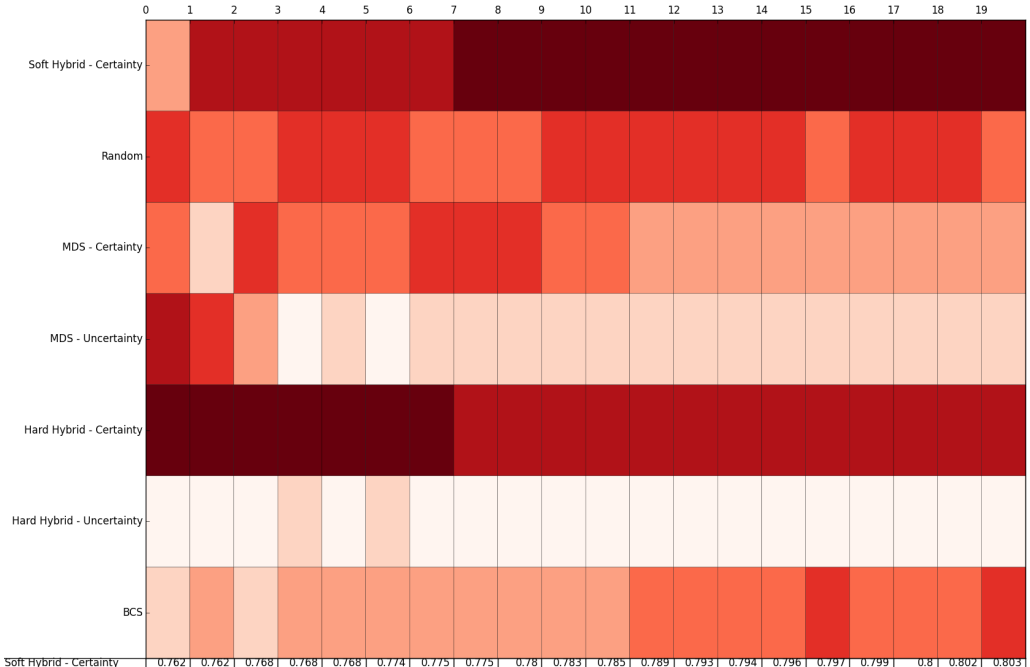
Soft Hybrid - Certainty	0.671	0.687	0.695	0.69	0.69	0.698	0.698	0.702	0.709	0.712	0.715	0.719	0.72	0.723	0.725	0.723	0.724	0.726	0.729	0.731
Random	0.668	0.672	0.675	0.673	0.677	0.682	0.687	0.692	0.695	0.698	0.702	0.707	0.711	0.716	0.72	0.725	0.727	0.73	0.734	0.736
MDS - Certainty	0.649	0.665	0.67	0.687	0.693	0.696	0.699	0.703	0.706	0.709	0.712	0.716	0.72	0.722	0.726	0.727	0.728	0.729	0.731	0.733
MDS - Uncertainty	0.653	0.662	0.671	0.68	0.688	0.695	0.691	0.692	0.697	0.704	0.707	0.707	0.707	0.711	0.716	0.718	0.721	0.724	0.727	0.729
Hard Hybrid - Certainty	0.698	0.703	0.695	0.695	0.689	0.689	0.691	0.697	0.7	0.701	0.703	0.705	0.708	0.711	0.714	0.715	0.719	0.721	0.722	0.723
Hard Hybrid - Uncertainty	0.683	0.698	0.684	0.687	0.698	0.695	0.693	0.695	0.696	0.7	0.7	0.704	0.711	0.714	0.719	0.72	0.725	0.728	0.732	0.734
BCS	0.687	0.681	0.694	0.703	0.704	0.706	0.711	0.717	0.722	0.727	0.728	0.73	0.733	0.735	0.737	0.737	0.738	0.74	0.742	0.742

LCP_PileEtFace_2011-11-19_085700, Head Track: 243, Speaker Track: 211, OCR Name: 40 HvH 173X173 SvS 162X162 SvH 289



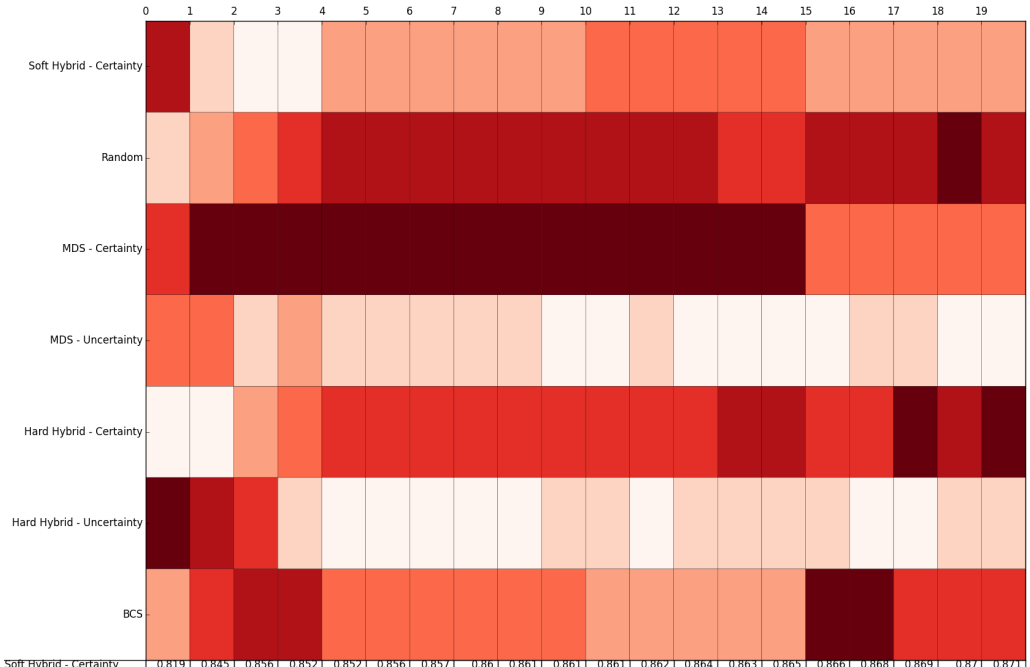
Soft Hybrid - Certainty	0.886	0.89	0.891	0.892	0.892	0.894	0.896	0.898	0.898	0.899	0.9	0.901	0.902	0.904	0.905	0.906	0.907	0.908	0.909	0.91
Random	0.873	0.885	0.888	0.891	0.891	0.891	0.894	0.896	0.898	0.899	0.901	0.903	0.905	0.907	0.908	0.91	0.911	0.912	0.914	0.915
MDS - Certainty	0.891	0.896	0.899	0.897	0.897	0.898	0.899	0.901	0.904	0.903	0.903	0.904	0.904	0.906	0.908	0.909	0.91	0.912	0.913	0.914
MDS - Uncertainty	0.875	0.89	0.895	0.898	0.901	0.905	0.906	0.906	0.907	0.909	0.91	0.911	0.913	0.914	0.915	0.916	0.917	0.918	0.919	0.919
Hard Hybrid - Certainty	0.89	0.881	0.88	0.876	0.877	0.872	0.877	0.88	0.883	0.887	0.89	0.892	0.893	0.894	0.896	0.898	0.899	0.901	0.902	0.902
Hard Hybrid - Uncertainty	0.894	0.894	0.888	0.893	0.889	0.887	0.89	0.893	0.897	0.899	0.9	0.902	0.904	0.906	0.908	0.909	0.911	0.911	0.912	0.913
BCS	0.873	0.888	0.889	0.885	0.882	0.882	0.882	0.882	0.885	0.885	0.887	0.888	0.89	0.891	0.893	0.895	0.896	0.898	0.899	0.9

LCP_PileEtFace_2011-12-01_192700, Head Track: 389, Speaker Track: 255, OCR Name: 51 HvH 250X250 SvS 197X197 SvH 349



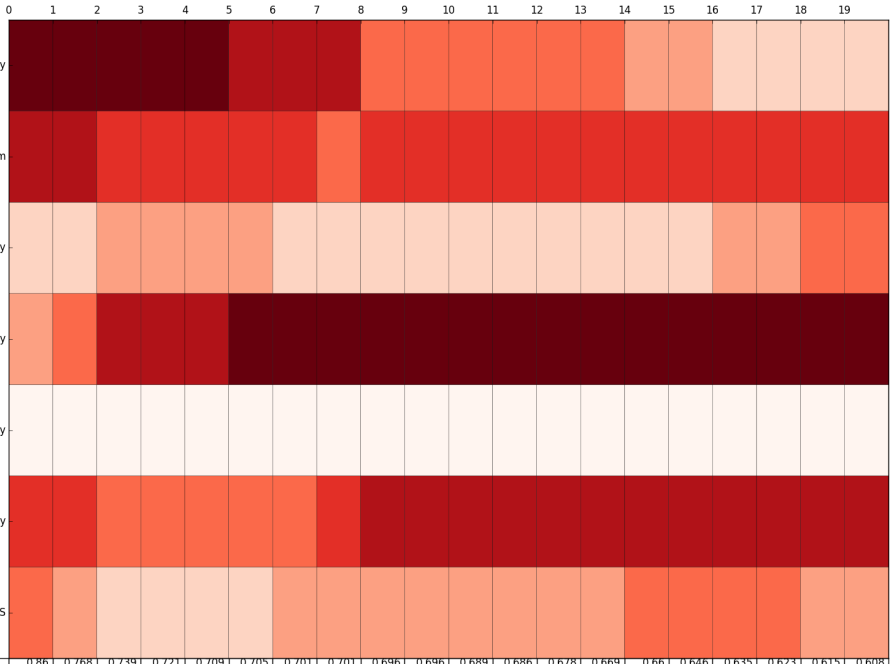
Soft Hybrid - Certainty	0.762	0.762	0.768	0.768	0.768	0.774	0.775	0.775	0.78	0.783	0.785	0.789	0.793	0.794	0.796	0.797	0.799	0.8	0.802	0.803
Random	0.755	0.768	0.776	0.772	0.776	0.781	0.786	0.79	0.793	0.795	0.796	0.8	0.803	0.805	0.806	0.808	0.808	0.808	0.809	0.811
MDS - Certainty	0.76	0.773	0.775	0.78	0.78	0.782	0.785	0.788	0.792	0.796	0.8	0.803	0.805	0.807	0.81	0.811	0.813	0.814	0.815	0.815
MDS - Uncertainty	0.744	0.767	0.777	0.792	0.794	0.799	0.799	0.802	0.805	0.809	0.811	0.813	0.813	0.815	0.817	0.818	0.819	0.82	0.82	0.821
Hard Hybrid - Certainty	0.707	0.746	0.76	0.76	0.767	0.767	0.771	0.777	0.782	0.785	0.789	0.793	0.795	0.797	0.798	0.799	0.801	0.803	0.804	0.805
Hard Hybrid - Uncertainty	0.785	0.783	0.788	0.787	0.795	0.798	0.803	0.808	0.812	0.814	0.815	0.817	0.819	0.82	0.822	0.823	0.824	0.824	0.825	0.825
BCS	0.782	0.772	0.781	0.783	0.789	0.793	0.795	0.798	0.798	0.8	0.802	0.803	0.805	0.806	0.807	0.808	0.81	0.81	0.81	0.811

LCP_PileEtFace_2012-01-12_192700, Head Track: 330, Speaker Track: 269, OCR Name: 52 HvH 228X228 SvS 218X218 SvH 373



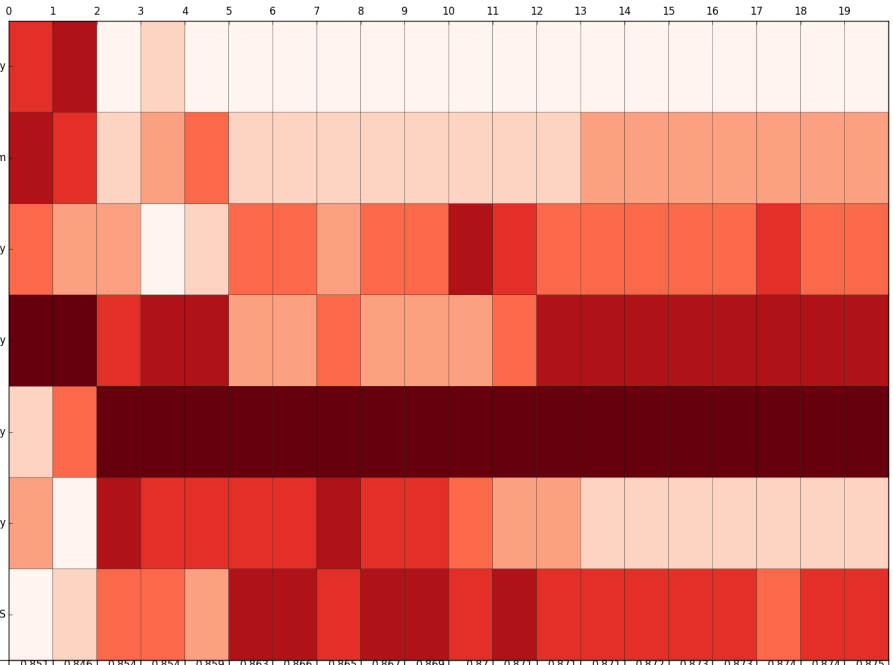
Soft Hybrid - Certainty	0.819	0.845	0.856	0.852	0.852	0.856	0.857	0.86	0.861	0.861	0.861	0.862	0.864	0.863	0.865	0.866	0.868	0.869	0.87	0.871
Random	0.864	0.843	0.847	0.847	0.849	0.845	0.847	0.849	0.852	0.855	0.857	0.858	0.861	0.863	0.865	0.865	0.866	0.866	0.867	0.868
MDS - Certainty	0.83	0.822	0.834	0.838	0.838	0.842	0.84	0.846	0.848	0.85	0.854	0.858	0.86	0.862	0.864	0.866	0.867	0.868	0.87	0.871
MDS - Uncertainty	0.842	0.842	0.849	0.848	0.853	0.856	0.859	0.863	0.866	0.869	0.87	0.87	0.872	0.873	0.875	0.876	0.875	0.875	0.876	0.877
Hard Hybrid - Certainty	0.864	0.849	0.847	0.848	0.849	0.849	0.851	0.854	0.855	0.857	0.857	0.859	0.861	0.862	0.864	0.865	0.866	0.865	0.867	0.867
Hard Hybrid - Uncertainty	0.807	0.839	0.844	0.852	0.855	0.859	0.862	0.866	0.868	0.869	0.87	0.871	0.872	0.873	0.875	0.876	0.876	0.877	0.875	0.876
BCS	0.843	0.84	0.844	0.847	0.85	0.854	0.856	0.859	0.86	0.86	0.862	0.864	0.865	0.866	0.867	0.865	0.866	0.867	0.868	0.869

LCP_PileEtFace_2012-01-19_192700, Head Track: 302, Speaker Track: 246, OCR Name: 43 HvH 222X222 SvS 191X191 SvH 290



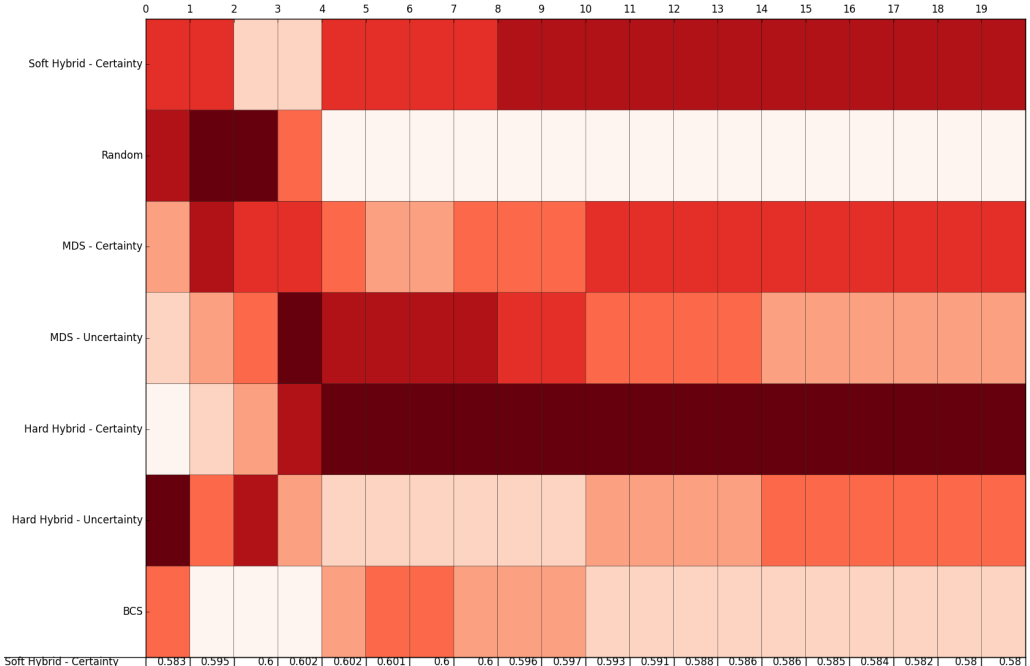
Soft Hybrid - Certainty	0.86	0.768	0.739	0.721	0.709	0.705	0.701	0.701	0.696	0.696	0.689	0.686	0.678	0.669	0.66	0.646	0.635	0.623	0.615	0.608
Random	0.863	0.823	0.81	0.789	0.758	0.737	0.724	0.711	0.695	0.688	0.681	0.668	0.65	0.636	0.624	0.614	0.608	0.596	0.587	0.579
MDS - Certainty	0.887	0.869	0.847	0.839	0.83	0.804	0.793	0.777	0.763	0.749	0.732	0.723	0.705	0.686	0.668	0.649	0.632	0.611	0.591	0.582
MDS - Uncertainty	0.88	0.835	0.794	0.759	0.713	0.689	0.66	0.634	0.619	0.599	0.589	0.58	0.573	0.559	0.549	0.543	0.539	0.528	0.519	0.509
Hard Hybrid - Certainty	0.897	0.884	0.883	0.876	0.872	0.854	0.838	0.816	0.792	0.767	0.756	0.738	0.721	0.706	0.692	0.674	0.66	0.644	0.635	0.625
Hard Hybrid - Uncertainty	0.873	0.827	0.818	0.809	0.802	0.768	0.738	0.71	0.689	0.663	0.643	0.621	0.598	0.584	0.571	0.563	0.556	0.544	0.539	0.532
BCS	0.88	0.86	0.856	0.856	0.846	0.819	0.792	0.775	0.755	0.74	0.724	0.707	0.688	0.672	0.652	0.637	0.62	0.609	0.599	0.593

LCP_PileEtFace_2012-01-26_192700, Head Track: 380, Speaker Track: 273, OCR Name: 50 HvH 223X223 SvS 207X207 SvH 349



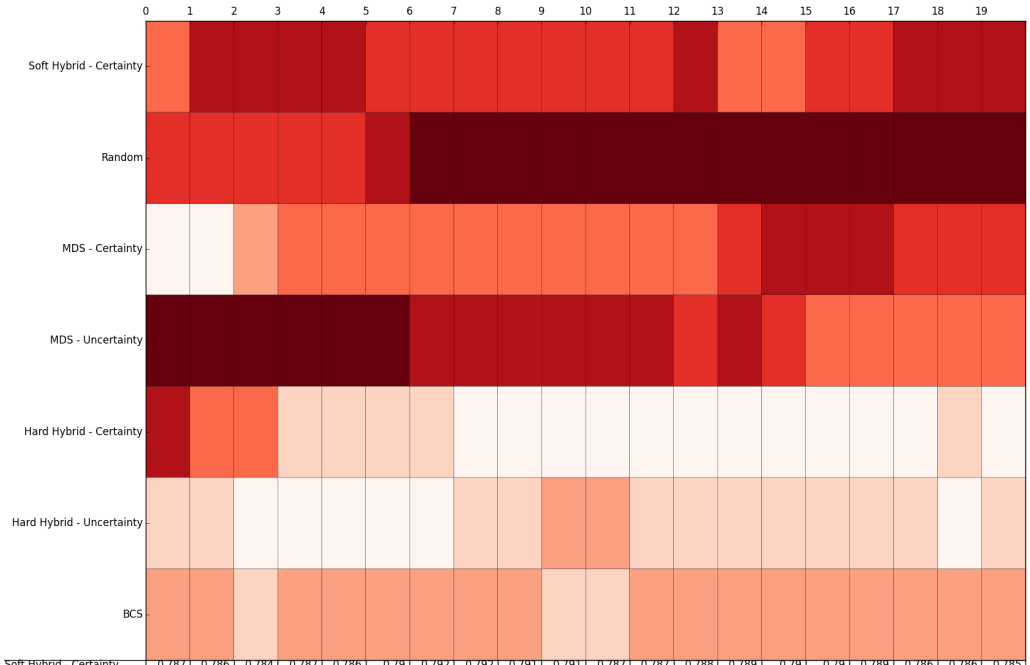
Soft Hybrid - Certainty	0.851	0.846	0.854	0.854	0.859	0.863	0.866	0.865	0.867	0.869	0.87	0.871	0.871	0.871	0.872	0.873	0.873	0.874	0.874	0.875
Random	0.85	0.849	0.854	0.853	0.85	0.853	0.856	0.856	0.857	0.859	0.86	0.862	0.865	0.864	0.866	0.867	0.867	0.868	0.869	0.87
MDS - Certainty	0.852	0.855	0.853	0.857	0.854	0.852	0.853	0.854	0.854	0.856	0.855	0.856	0.858	0.86	0.861	0.863	0.865	0.866	0.867	0.868
MDS - Uncertainty	0.826	0.834	0.844	0.844	0.847	0.852	0.853	0.853	0.856	0.856	0.857	0.857	0.856	0.857	0.859	0.86	0.862	0.864	0.865	0.866
Hard Hybrid - Certainty	0.854	0.851	0.841	0.839	0.842	0.843	0.84	0.842	0.846	0.846	0.849	0.85	0.852	0.853	0.856	0.858	0.86	0.862	0.863	0.864
Hard Hybrid - Uncertainty	0.852	0.862	0.842	0.849	0.848	0.85	0.852	0.849	0.853	0.855	0.857	0.859	0.861	0.864	0.866	0.869	0.869	0.871	0.872	0.873
BCS	0.862	0.857	0.848	0.85	0.852	0.85	0.851	0.851	0.853	0.855	0.856	0.856	0.858	0.859	0.861	0.863	0.865	0.867	0.867	0.867

LCP_TopQuestions_2012-01-25_232700, Head Track: 364, Speaker Track: 134, OCR Name: 11 HvH 204X204 SvS 102X102 SvH 310



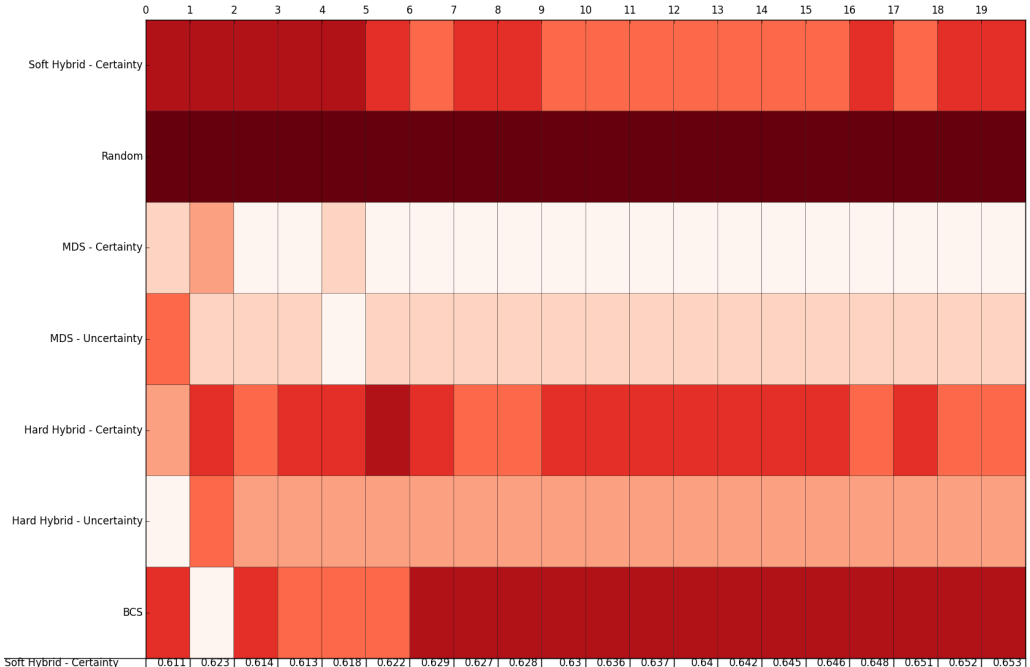
Soft Hybrid - Certainty	0.583	0.595	0.6	0.602	0.602	0.601	0.6	0.6	0.596	0.597	0.593	0.591	0.588	0.586	0.586	0.585	0.584	0.582	0.58	0.58
Random	0.583	0.579	0.589	0.599	0.611	0.619	0.624	0.624	0.627	0.632	0.636	0.638	0.638	0.64	0.643	0.645	0.646	0.647	0.648	0.651
MDS - Certainty	0.601	0.594	0.592	0.599	0.605	0.611	0.608	0.604	0.604	0.601	0.599	0.597	0.595	0.595	0.593	0.591	0.591	0.591	0.59	0.591
MDS - Uncertainty	0.608	0.598	0.594	0.593	0.594	0.593	0.597	0.595	0.598	0.599	0.6	0.6	0.6	0.601	0.603	0.605	0.606	0.607	0.607	0.608
Hard Hybrid - Certainty	0.615	0.604	0.595	0.594	0.589	0.586	0.585	0.581	0.575	0.571	0.569	0.57	0.57	0.571	0.569	0.567	0.57	0.569	0.57	0.57
Hard Hybrid - Uncertainty	0.58	0.596	0.591	0.6	0.608	0.613	0.617	0.615	0.617	0.617	0.612	0.608	0.605	0.603	0.602	0.6	0.597	0.595	0.594	0.593
BCS	0.601	0.607	0.617	0.613	0.606	0.605	0.604	0.609	0.611	0.611	0.614	0.613	0.612	0.611	0.61	0.612	0.614	0.615	0.615	0.616

LCP_TopQuestions_2012-02-14_232200, Head Track: 249, Speaker Track: 155, OCR Name: 13 HvH 133X133 SvS 127X127 SvH 272



Soft Hybrid - Certainty	0.787	0.786	0.784	0.787	0.786	0.79	0.792	0.792	0.791	0.791	0.787	0.787	0.788	0.789	0.79	0.79	0.789	0.786	0.786	0.785
Random	0.779	0.787	0.796	0.79	0.796	0.788	0.788	0.783	0.778	0.778	0.777	0.776	0.776	0.775	0.774	0.776	0.777	0.777	0.776	0.777
MDS - Certainty	0.828	0.816	0.807	0.803	0.802	0.801	0.8	0.802	0.796	0.795	0.792	0.792	0.791	0.789	0.787	0.787	0.787	0.789	0.79	0.791
MDS - Uncertainty	0.769	0.762	0.768	0.773	0.782	0.787	0.79	0.788	0.786	0.784	0.786	0.785	0.788	0.788	0.789	0.79	0.791	0.792	0.794	0.795
Hard Hybrid - Certainty	0.771	0.795	0.803	0.809	0.811	0.813	0.814	0.814	0.814	0.815	0.814	0.812	0.81	0.808	0.807	0.805	0.801	0.802	0.801	0.801
Hard Hybrid - Uncertainty	0.824	0.814	0.815	0.815	0.818	0.817	0.816	0.812	0.811	0.807	0.805	0.807	0.806	0.805	0.803	0.803	0.799	0.8	0.802	0.8
BCS	0.796	0.806	0.811	0.808	0.81	0.811	0.813	0.809	0.809	0.808	0.808	0.805	0.803	0.802	0.801	0.799	0.796	0.797	0.797	0.798

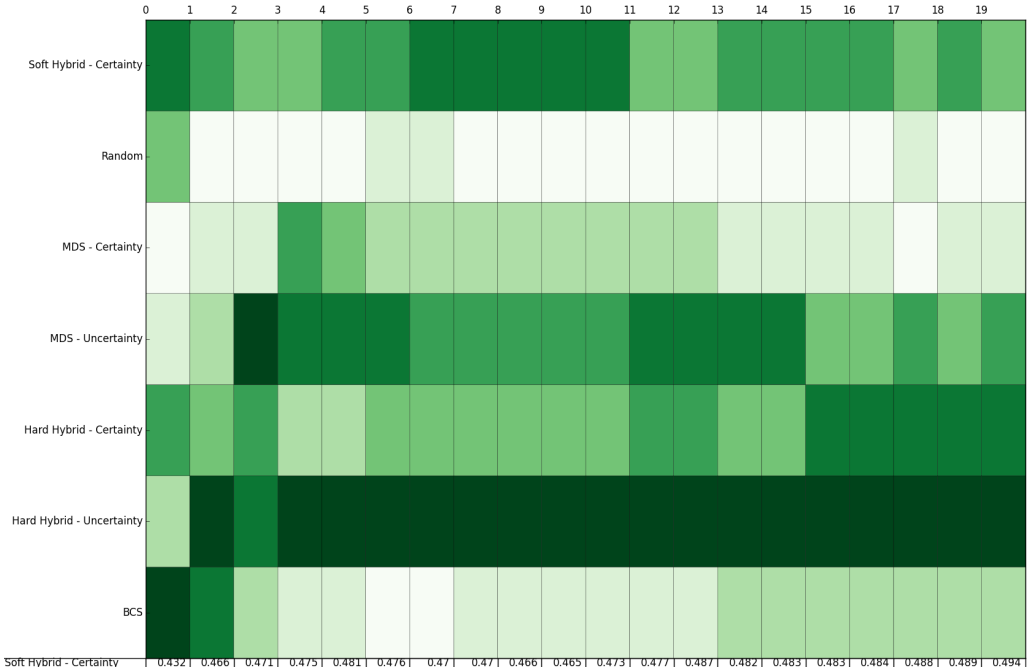
LCP_TopQuestions_2012-02-22_232700, Head Track: 308, Speaker Track: 158, OCR Name: 15 HvH 170X170 SvS 129X129 SvH 272



Soft Hybrid - Certainty	0.611	0.623	0.614	0.613	0.618	0.622	0.629	0.627	0.628	0.63	0.636	0.637	0.64	0.642	0.645	0.646	0.648	0.651	0.652	0.653
Random	0.573	0.562	0.551	0.567	0.571	0.573	0.579	0.589	0.598	0.602	0.607	0.61	0.615	0.619	0.62	0.625	0.627	0.629	0.631	0.635
MDS - Certainty	0.67	0.654	0.654	0.66	0.651	0.66	0.668	0.665	0.664	0.663	0.666	0.668	0.672	0.675	0.677	0.679	0.682	0.683	0.686	0.688
MDS - Uncertainty	0.628	0.654	0.651	0.651	0.651	0.656	0.652	0.653	0.655	0.657	0.658	0.662	0.662	0.663	0.667	0.668	0.671	0.675	0.677	0.679
Hard Hybrid - Certainty	0.633	0.629	0.629	0.619	0.619	0.62	0.621	0.628	0.63	0.63	0.63	0.633	0.637	0.64	0.643	0.646	0.649	0.651	0.653	0.655
Hard Hybrid - Uncertainty	0.676	0.631	0.634	0.637	0.639	0.64	0.647	0.651	0.649	0.644	0.647	0.648	0.653	0.656	0.658	0.66	0.662	0.663	0.666	0.669
BCS	0.625	0.655	0.627	0.624	0.627	0.624	0.618	0.62	0.622	0.619	0.622	0.623	0.624	0.626	0.628	0.63	0.631	0.635	0.638	0.64

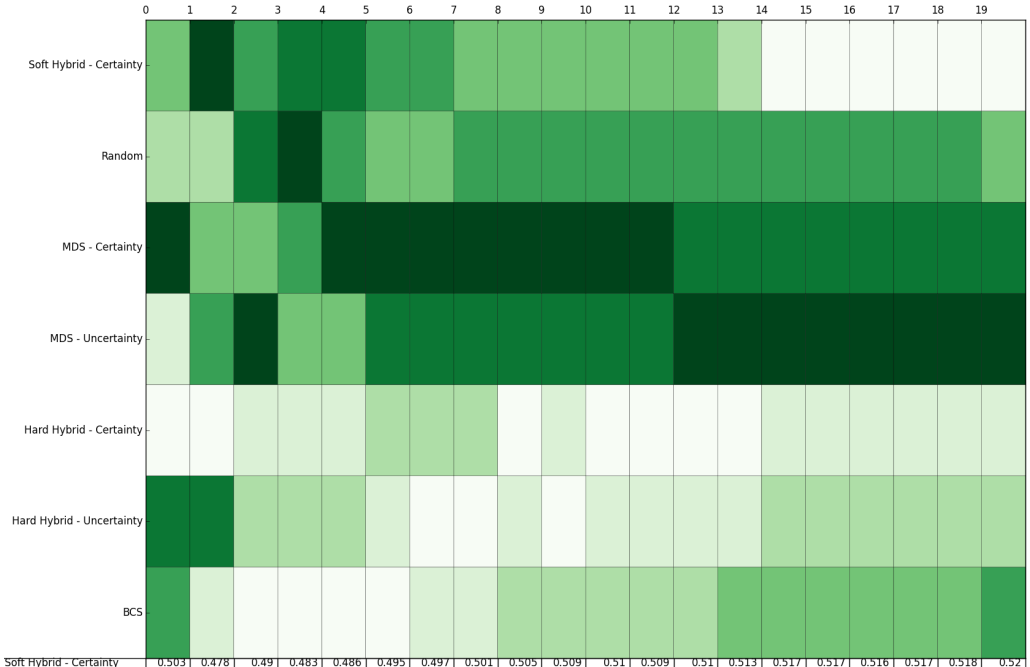
Speaker score using face annotation is given green in heat maps.

BFMTV_BFMStory_2012-01-10_175800, Head Track: 759, Speaker Track: 372, OCR Name: 36 HvH 401X401 SvS 291X291 SvH 737



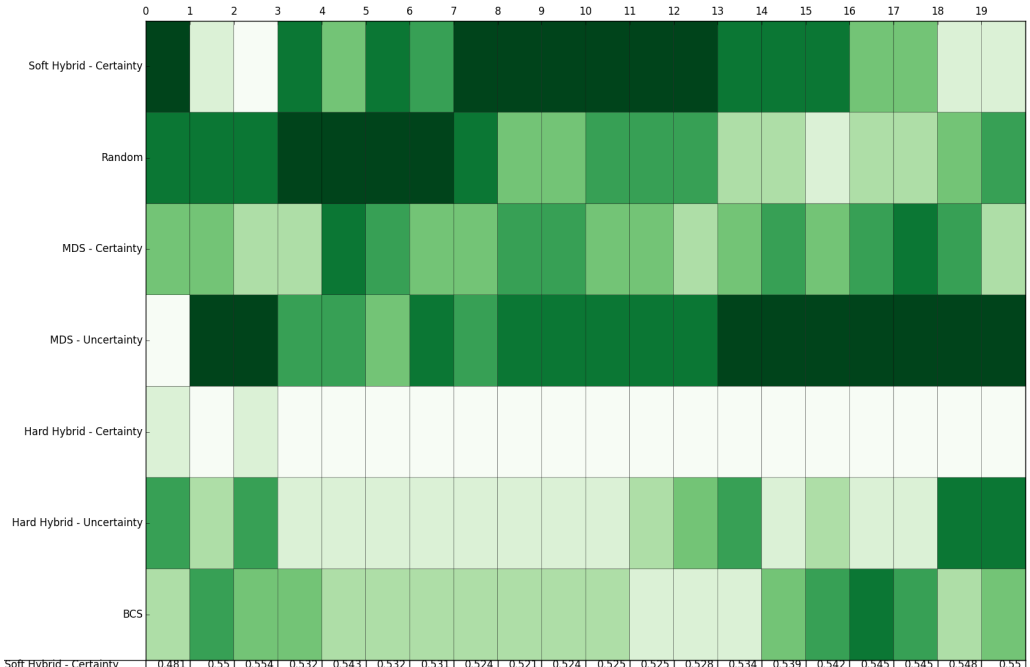
Soft Hybrid - Certainty	0.432	0.466	0.471	0.475	0.481	0.476	0.47	0.47	0.466	0.465	0.473	0.477	0.487	0.482	0.483	0.483	0.484	0.488	0.489	0.494
Random	0.441	0.492	0.509	0.505	0.51	0.502	0.51	0.51	0.512	0.515	0.509	0.509	0.51	0.51	0.512	0.513	0.513	0.514	0.517	0.519
MDS - Certainty	0.5	0.49	0.488	0.471	0.485	0.498	0.501	0.504	0.499	0.503	0.497	0.497	0.503	0.507	0.505	0.51	0.51	0.516	0.514	0.518
MDS - Uncertainty	0.457	0.475	0.449	0.453	0.442	0.459	0.479	0.478	0.467	0.468	0.474	0.474	0.473	0.474	0.479	0.484	0.484	0.484	0.49	0.49
Hard Hybrid - Certainty	0.439	0.468	0.466	0.478	0.491	0.486	0.49	0.488	0.487	0.484	0.476	0.477	0.481	0.483	0.484	0.482	0.48	0.481	0.486	0.487
Hard Hybrid - Uncertainty	0.448	0.442	0.457	0.447	0.442	0.432	0.447	0.442	0.449	0.456	0.46	0.464	0.464	0.469	0.471	0.472	0.474	0.474	0.475	0.476
BCS	0.393	0.449	0.487	0.494	0.499	0.514	0.519	0.507	0.507	0.507	0.506	0.508	0.507	0.504	0.503	0.497	0.499	0.503	0.501	0.501

BFMTV_BFMStory_2012-01-23_175800, Head Track: 782, Speaker Track: 338, OCR Name: 34 HvH 406X406 SvS 247X247 SvH 556



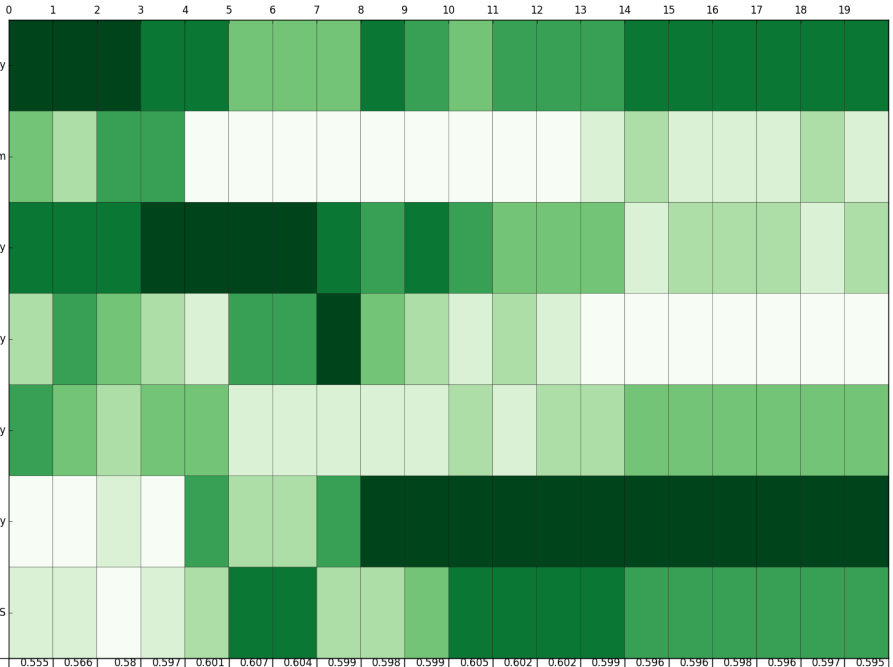
Soft Hybrid - Certainty	0.503	0.478	0.49	0.483	0.486	0.495	0.497	0.501	0.505	0.509	0.51	0.509	0.51	0.513	0.517	0.517	0.516	0.517	0.518	0.52
Random	0.508	0.503	0.49	0.48	0.488	0.498	0.499	0.498	0.501	0.505	0.5	0.501	0.504	0.509	0.509	0.506	0.507	0.508	0.512	0.514
MDS - Certainty	0.456	0.501	0.501	0.488	0.476	0.484	0.482	0.486	0.487	0.489	0.492	0.491	0.496	0.496	0.495	0.499	0.5	0.502	0.505	0.504
MDS - Uncertainty	0.536	0.5	0.489	0.5	0.495	0.493	0.495	0.494	0.495	0.493	0.496	0.495	0.492	0.488	0.487	0.488	0.489	0.489	0.49	0.49
Hard Hybrid - Certainty	0.561	0.543	0.517	0.516	0.513	0.511	0.51	0.509	0.518	0.519	0.519	0.518	0.518	0.519	0.517	0.515	0.516	0.517	0.517	0.519
Hard Hybrid - Uncertainty	0.467	0.495	0.501	0.502	0.512	0.514	0.517	0.517	0.518	0.52	0.516	0.517	0.518	0.517	0.517	0.515	0.515	0.515	0.517	0.518
BCS	0.495	0.529	0.522	0.521	0.515	0.516	0.515	0.512	0.514	0.515	0.512	0.511	0.513	0.512	0.513	0.515	0.514	0.515	0.514	0.513

BFMTV_BFMStory_2012-02-14_175800, Head Track: 753, Speaker Track: 348, OCR Name: 33 HvH 405X405 SvS 272X272 SvH 643



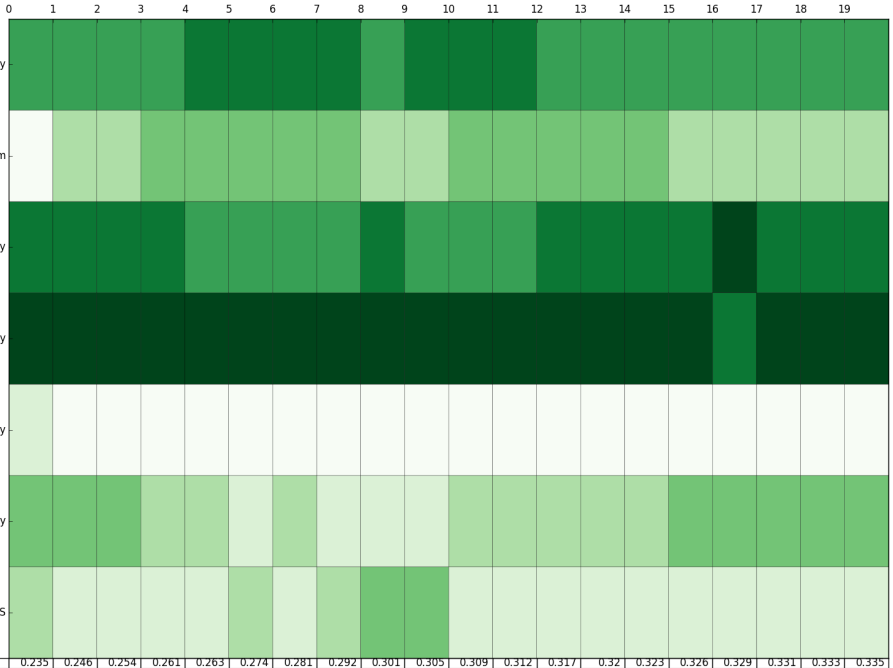
Soft Hybrid - Certainty	0.481	0.55	0.554	0.532	0.543	0.532	0.531	0.524	0.521	0.524	0.525	0.525	0.528	0.534	0.539	0.542	0.545	0.545	0.548	0.55
Random	0.484	0.519	0.526	0.523	0.519	0.511	0.519	0.526	0.534	0.538	0.537	0.538	0.539	0.544	0.546	0.547	0.548	0.547	0.546	0.545
MDS - Certainty	0.513	0.536	0.547	0.542	0.527	0.535	0.532	0.536	0.533	0.537	0.54	0.544	0.546	0.542	0.543	0.546	0.545	0.544	0.546	0.547
MDS - Uncertainty	0.569	0.516	0.518	0.534	0.542	0.536	0.529	0.53	0.522	0.525	0.525	0.528	0.528	0.533	0.533	0.533	0.53	0.532	0.536	0.535
Hard Hybrid - Certainty	0.554	0.559	0.55	0.561	0.567	0.571	0.573	0.574	0.574	0.572	0.572	0.568	0.571	0.572	0.57	0.574	0.572	0.575	0.576	0.575
Hard Hybrid - Uncertainty	0.503	0.539	0.535	0.553	0.562	0.562	0.568	0.562	0.566	0.554	0.551	0.549	0.544	0.541	0.546	0.546	0.55	0.548	0.543	0.544
BCS	0.545	0.524	0.54	0.538	0.545	0.542	0.545	0.551	0.55	0.553	0.547	0.553	0.553	0.547	0.546	0.545	0.545	0.545	0.547	0.547

BFMTV_BFMStory_2012-02-20_175800, Head Track: 871, Speaker Track: 342, OCR Name: 30 HvH 450X450 SvS 267X267 SvH 710



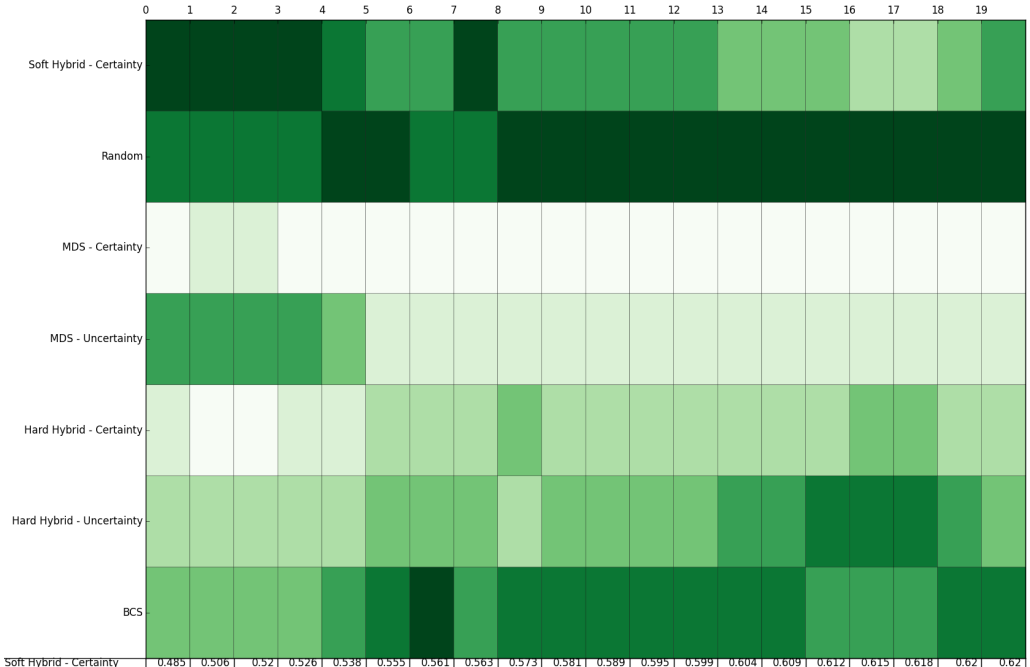
Soft Hybrid - Certainty	0.555	0.566	0.58	0.597	0.601	0.607	0.604	0.599	0.598	0.599	0.605	0.602	0.602	0.599	0.596	0.596	0.598	0.596	0.597	0.595
Random	0.626	0.609	0.598	0.6	0.618	0.616	0.616	0.619	0.619	0.617	0.617	0.615	0.614	0.613	0.61	0.61	0.612	0.61	0.607	0.608
MDS - Certainty	0.599	0.597	0.588	0.581	0.582	0.594	0.595	0.599	0.6	0.599	0.6	0.603	0.605	0.607	0.611	0.608	0.606	0.608	0.611	0.608
MDS - Uncertainty	0.643	0.601	0.614	0.612	0.612	0.606	0.603	0.597	0.602	0.608	0.606	0.607	0.609	0.614	0.617	0.618	0.616	0.619	0.62	0.621
Hard Hybrid - Certainty	0.621	0.602	0.619	0.608	0.61	0.615	0.614	0.613	0.609	0.609	0.606	0.608	0.606	0.608	0.608	0.606	0.605	0.603	0.603	0.605
Hard Hybrid - Uncertainty	0.653	0.629	0.629	0.62	0.607	0.611	0.61	0.599	0.594	0.593	0.593	0.591	0.589	0.59	0.586	0.588	0.591	0.592	0.59	0.589
BCS	0.645	0.621	0.637	0.618	0.611	0.605	0.599	0.601	0.603	0.601	0.598	0.598	0.597	0.597	0.6	0.602	0.601	0.602	0.602	0.605

BFMTV_CultureEtVous_2012-01-13_065040, Head Track: 69, Speaker Track: 38, OCR Name: 3 HvH 29X29 SvS 28X28 SvH 42



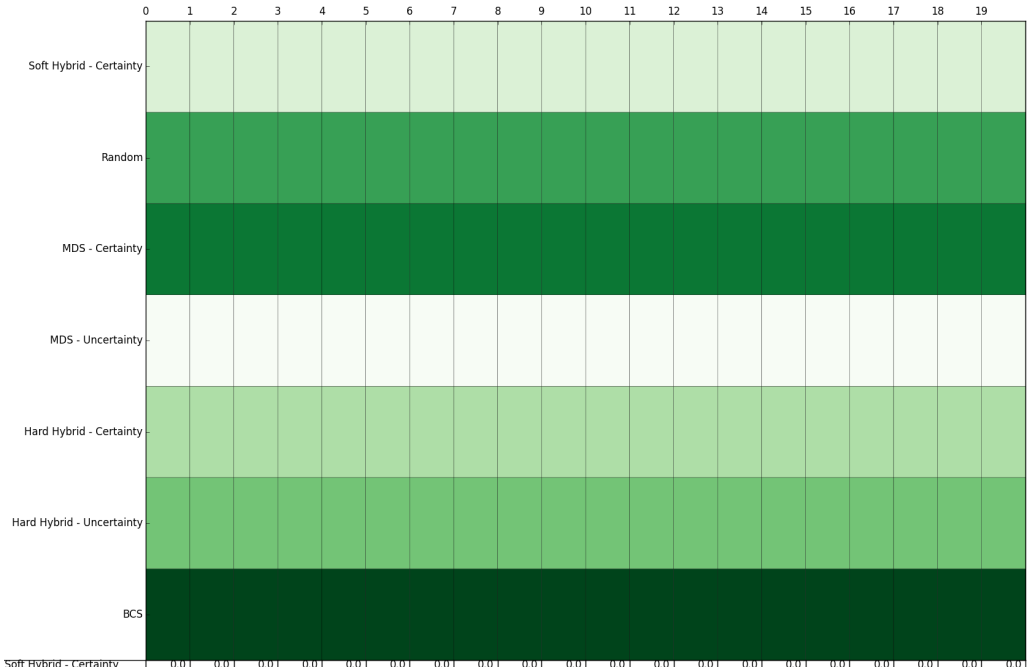
Soft Hybrid - Certainty	0.235	0.246	0.254	0.261	0.263	0.274	0.281	0.292	0.301	0.305	0.309	0.312	0.317	0.32	0.323	0.326	0.329	0.331	0.333	0.335
Random	0.25	0.269	0.289	0.299	0.305	0.315	0.321	0.329	0.335	0.338	0.339	0.336	0.338	0.338	0.34	0.344	0.344	0.346	0.346	0.347
MDS - Certainty	0.227	0.238	0.247	0.259	0.271	0.28	0.287	0.293	0.299	0.307	0.31	0.313	0.314	0.314	0.316	0.317	0.317	0.319	0.321	0.322
MDS - Uncertainty	0.215	0.226	0.233	0.245	0.255	0.266	0.276	0.282	0.289	0.294	0.298	0.304	0.307	0.309	0.312	0.315	0.317	0.318	0.319	0.32
Hard Hybrid - Certainty	0.249	0.278	0.31	0.328	0.33	0.331	0.339	0.342	0.344	0.344	0.344	0.346	0.347	0.348	0.35	0.35	0.351	0.352	0.352	0.353
Hard Hybrid - Uncertainty	0.238	0.257	0.283	0.299	0.31	0.319	0.326	0.333	0.337	0.34	0.339	0.339	0.34	0.34	0.341	0.342	0.343	0.343	0.345	0.345
BCS	0.239	0.276	0.29	0.304	0.313	0.319	0.327	0.332	0.334	0.338	0.341	0.342	0.344	0.345	0.346	0.345	0.347	0.347	0.348	0.35

BFMTV_CultureEtVous_2012-01-16_075000, Head Track: 163, Speaker Track: 42, OCR Name: 5 HvH 78X78 SvS 32X32 SvH 117



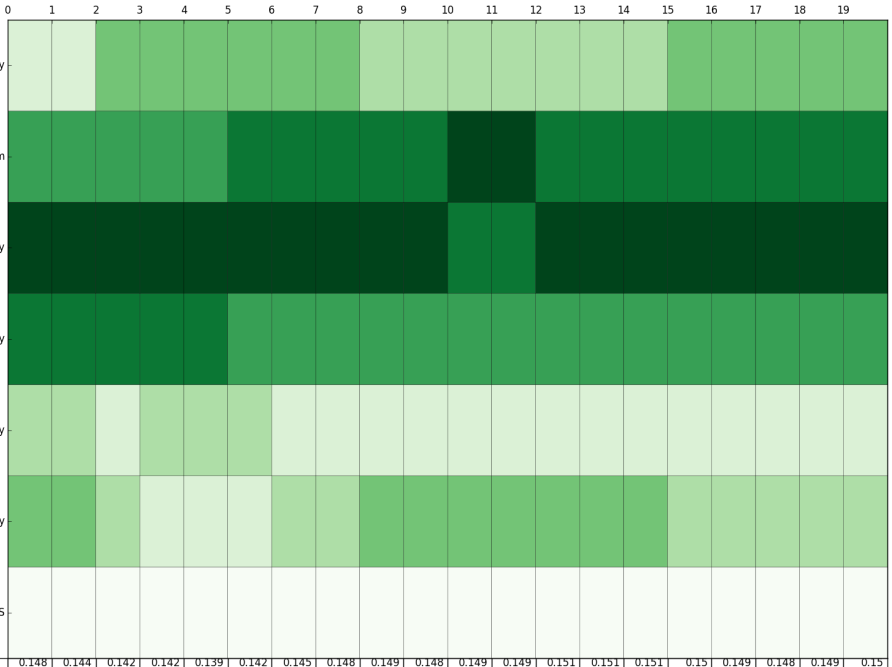
Soft Hybrid - Certainty	0.485	0.506	0.52	0.526	0.538	0.555	0.561	0.563	0.573	0.581	0.589	0.595	0.599	0.604	0.609	0.612	0.615	0.618	0.62	0.62
Random	0.49	0.516	0.521	0.533	0.532	0.538	0.558	0.565	0.57	0.574	0.58	0.586	0.589	0.592	0.595	0.597	0.602	0.608	0.612	0.615
MDS - Certainty	0.544	0.548	0.553	0.567	0.577	0.584	0.592	0.596	0.603	0.606	0.614	0.616	0.618	0.621	0.625	0.623	0.627	0.626	0.628	0.631
MDS - Uncertainty	0.508	0.523	0.53	0.542	0.552	0.567	0.576	0.585	0.59	0.597	0.604	0.611	0.613	0.614	0.618	0.62	0.624	0.624	0.627	0.626
Hard Hybrid - Certainty	0.541	0.548	0.562	0.562	0.566	0.564	0.565	0.572	0.577	0.586	0.593	0.597	0.601	0.608	0.611	0.614	0.614	0.617	0.621	0.623
Hard Hybrid - Uncertainty	0.525	0.538	0.544	0.553	0.559	0.563	0.564	0.57	0.58	0.583	0.59	0.597	0.6	0.604	0.606	0.609	0.612	0.612	0.618	0.621
BCS	0.514	0.535	0.54	0.55	0.548	0.554	0.557	0.566	0.572	0.578	0.585	0.591	0.595	0.599	0.605	0.611	0.613	0.613	0.618	0.62

BFMTV_CultureEtVous_2012-01-17_065040, Head Track: 48, Speaker Track: 49, OCR Name: 5 HvH 23X23 SvS 35X35 SvH 30



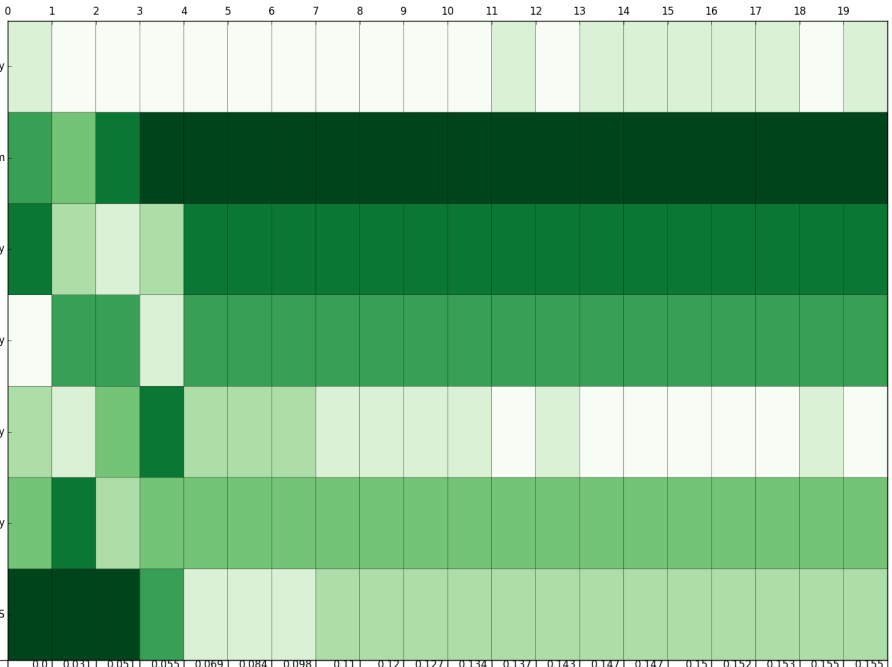
Soft Hybrid - Certainty	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Random	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
MDS - Certainty	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
MDS - Uncertainty	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Hard Hybrid - Certainty	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Hard Hybrid - Uncertainty	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
BCS	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

BFMTV_CultureEtVous_2012-01-18_065040, Head Track: 102, Speaker Track: 47, OCR Name: 6 HvH 41X41 SvS 37X37 SvH 42



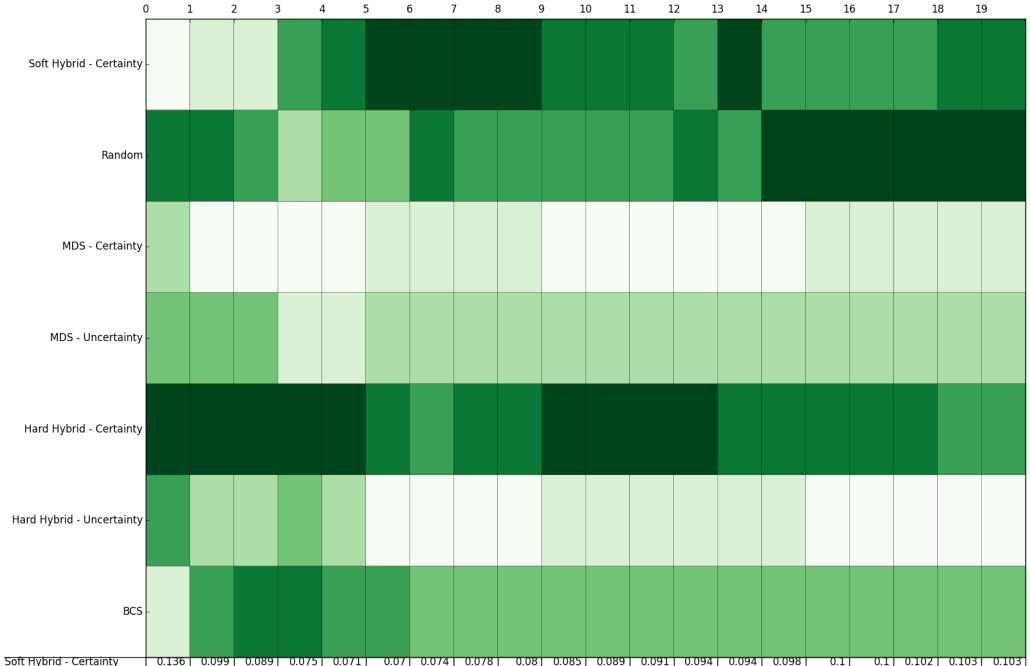
Soft Hybrid - Certainty	0.148	0.144	0.142	0.142	0.139	0.142	0.145	0.148	0.149	0.148	0.149	0.149	0.151	0.151	0.151	0.15	0.149	0.148	0.149	0.15
Random	0.132	0.133	0.135	0.133	0.134	0.134	0.131	0.131	0.132	0.132	0.131	0.132	0.133	0.135	0.135	0.135	0.136	0.135	0.135	0.135
MDS - Certainty	0.115	0.116	0.126	0.123	0.123	0.124	0.122	0.125	0.13	0.132	0.133	0.133	0.133	0.134	0.133	0.133	0.133	0.133	0.133	0.132
MDS - Uncertainty	0.124	0.125	0.126	0.127	0.131	0.134	0.138	0.137	0.138	0.138	0.137	0.137	0.137	0.14	0.14	0.14	0.141	0.141	0.141	0.141
Hard Hybrid - Certainty	0.146	0.142	0.153	0.15	0.151	0.149	0.151	0.152	0.154	0.154	0.153	0.153	0.154	0.153	0.153	0.153	0.153	0.153	0.152	0.151
Hard Hybrid - Uncertainty	0.14	0.139	0.144	0.152	0.153	0.153	0.15	0.149	0.148	0.146	0.146	0.146	0.146	0.149	0.15	0.151	0.15	0.151	0.152	0.151
BCS	0.151	0.148	0.155	0.162	0.162	0.161	0.162	0.161	0.158	0.161	0.161	0.163	0.161	0.159	0.159	0.159	0.161	0.16	0.16	0.161

BFMTV_CultureEtVous_2012-01-19_065040, Head Track: 78, Speaker Track: 41, OCR Name: 4 HvH 43X43 SvS 31X31 SvH 52



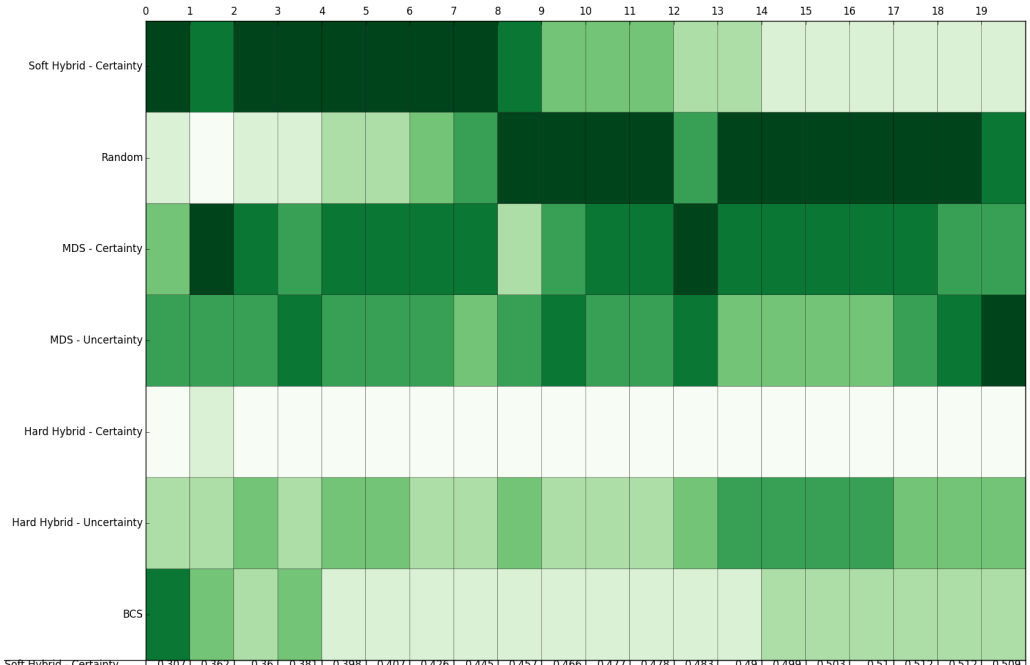
Soft Hybrid - Certainty	0.0	0.031	0.051	0.055	0.069	0.084	0.098	0.11	0.12	0.127	0.134	0.137	0.143	0.147	0.147	0.15	0.152	0.153	0.155	0.155
Random	0.0	0.011	0.015	0.015	0.036	0.041	0.05	0.055	0.061	0.069	0.072	0.074	0.076	0.078	0.08	0.081	0.083	0.085	0.085	0.086
MDS - Certainty	0.0	0.014	0.027	0.031	0.037	0.047	0.059	0.069	0.079	0.088	0.096	0.101	0.105	0.11	0.109	0.112	0.114	0.116	0.117	0.119
MDS - Uncertainty	0.0	0.0	0.017	0.032	0.039	0.053	0.068	0.073	0.084	0.088	0.096	0.103	0.109	0.112	0.115	0.12	0.12	0.122	0.124	0.125
Hard Hybrid - Certainty	0.0	0.014	0.018	0.023	0.049	0.072	0.092	0.105	0.115	0.123	0.132	0.139	0.143	0.15	0.154	0.157	0.156	0.156	0.155	0.158
Hard Hybrid - Uncertainty	0.0	0.0	0.019	0.03	0.044	0.069	0.082	0.093	0.098	0.102	0.107	0.111	0.117	0.121	0.124	0.127	0.13	0.131	0.134	0.137
BCS	0.0	0.0	0.008	0.027	0.056	0.081	0.097	0.103	0.111	0.119	0.123	0.131	0.133	0.135	0.14	0.142	0.144	0.148	0.15	0.151

BFMTV_CultureEtVous_2012-02-14_065040, Head Track: 104, Speaker Track: 48, OCR Name: 9 HvH 48X48 SvS 39X39 SvH 61



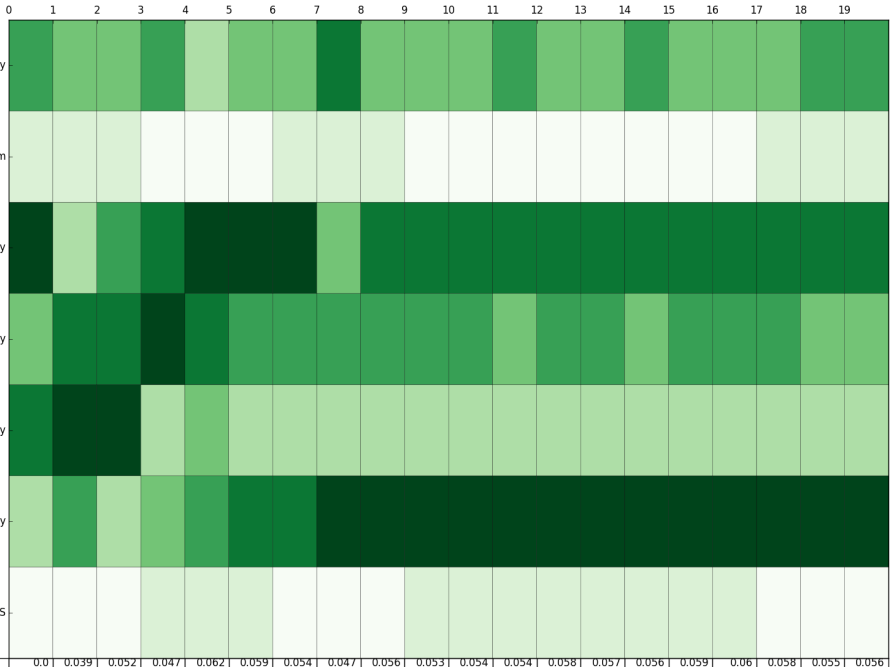
Soft Hybrid - Certainty	0.136	0.099	0.089	0.075	0.071	0.07	0.074	0.078	0.08	0.085	0.089	0.091	0.094	0.094	0.098	0.1	0.1	0.102	0.103	0.103
Random	0.046	0.072	0.072	0.082	0.091	0.083	0.082	0.091	0.088	0.089	0.092	0.092	0.094	0.097	0.096	0.096	0.094	0.095	0.097	0.093
MDS - Certainty	0.09	0.105	0.095	0.099	0.104	0.105	0.111	0.11	0.114	0.122	0.121	0.122	0.121	0.121	0.123	0.121	0.122	0.122	0.122	0.122
MDS - Uncertainty	0.079	0.088	0.074	0.087	0.092	0.103	0.104	0.107	0.11	0.111	0.112	0.114	0.115	0.117	0.116	0.118	0.119	0.119	0.12	0.12
Hard Hybrid - Certainty	0.044	0.046	0.042	0.048	0.06	0.075	0.082	0.088	0.086	0.084	0.084	0.087	0.091	0.095	0.098	0.099	0.1	0.101	0.104	0.104
Hard Hybrid - Uncertainty	0.055	0.099	0.085	0.079	0.091	0.108	0.112	0.117	0.121	0.118	0.119	0.118	0.12	0.12	0.121	0.123	0.124	0.126	0.129	0.128
BCS	0.115	0.082	0.071	0.075	0.078	0.083	0.091	0.098	0.098	0.102	0.105	0.107	0.108	0.109	0.107	0.11	0.112	0.111	0.111	0.112

BFMTV_CultureEtVous_2012-02-15_065040, Head Track: 180, Speaker Track: 34, OCR Name: 4 HvH 67X67 SvS 27X27 SvH 113



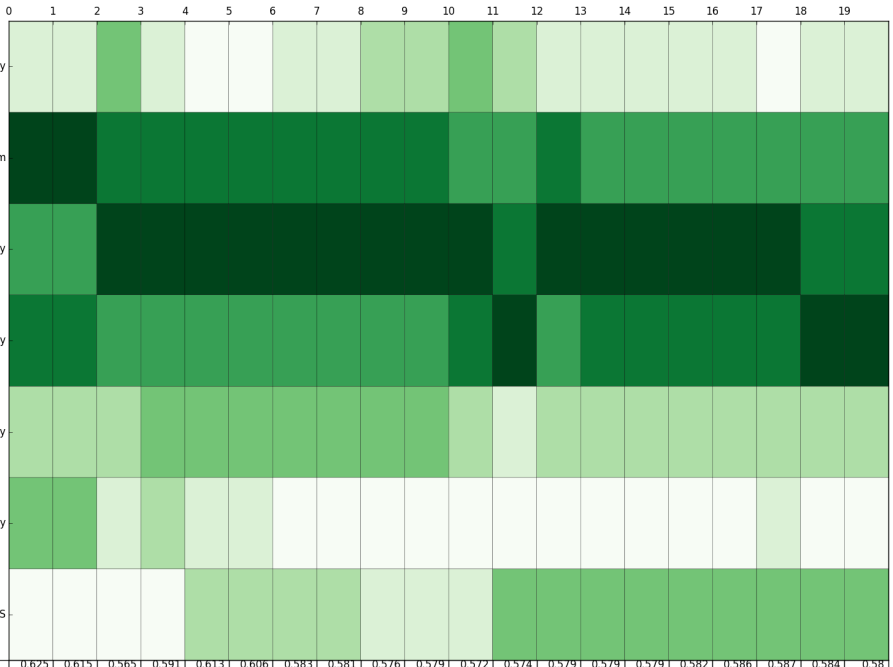
Soft Hybrid - Certainty	0.307	0.362	0.36	0.381	0.398	0.407	0.426	0.445	0.457	0.466	0.477	0.478	0.483	0.49	0.499	0.503	0.51	0.512	0.512	0.509
Random	0.413	0.423	0.421	0.439	0.44	0.447	0.448	0.455	0.45	0.454	0.462	0.468	0.475	0.473	0.475	0.474	0.476	0.479	0.478	0.48
MDS - Certainty	0.362	0.313	0.374	0.398	0.406	0.412	0.435	0.454	0.463	0.463	0.468	0.47	0.472	0.478	0.477	0.478	0.481	0.48	0.48	0.482
MDS - Uncertainty	0.355	0.369	0.384	0.396	0.407	0.43	0.441	0.456	0.457	0.459	0.471	0.471	0.474	0.482	0.485	0.487	0.484	0.482	0.479	0.479
Hard Hybrid - Certainty	0.416	0.408	0.435	0.445	0.471	0.486	0.49	0.497	0.494	0.494	0.505	0.504	0.51	0.515	0.518	0.516	0.52	0.516	0.519	0.518
Hard Hybrid - Uncertainty	0.378	0.401	0.409	0.435	0.433	0.442	0.463	0.462	0.462	0.469	0.48	0.482	0.48	0.482	0.48	0.486	0.484	0.492	0.496	0.498
BCS	0.339	0.376	0.415	0.431	0.456	0.465	0.473	0.468	0.48	0.485	0.492	0.494	0.497	0.495	0.492	0.495	0.498	0.499	0.5	0.508

LCP_CaVousRegarde_2011-12-20_204600, Head Track: 397, Speaker Track: 332, OCR Name: 30 HvH 282X282 SvS 268X268 SvH 406



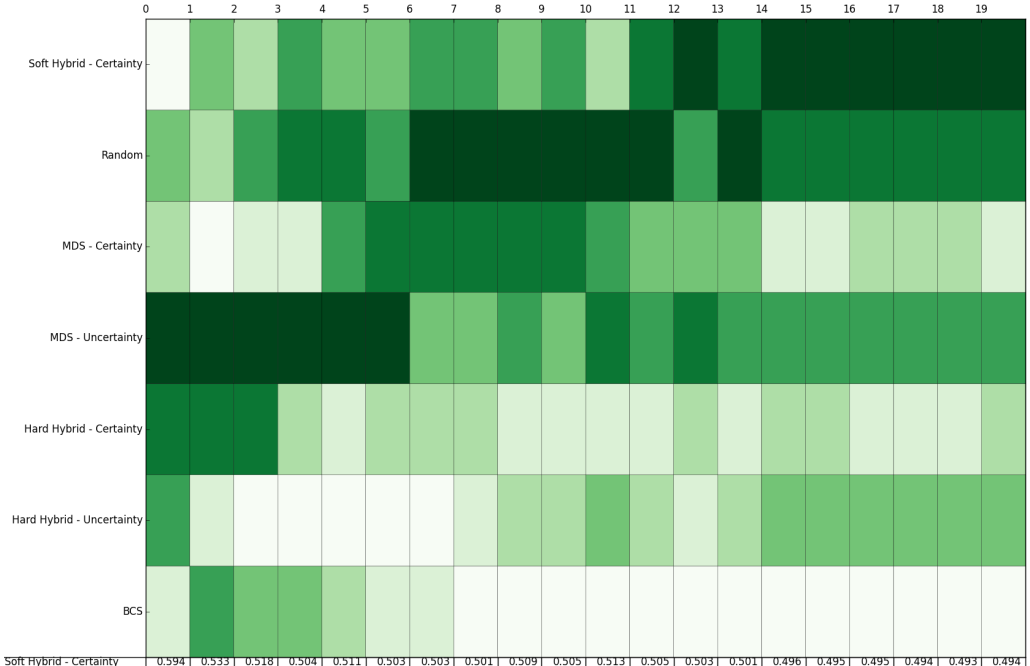
Soft Hybrid - Certainty	0.0	0.039	0.052	0.047	0.062	0.059	0.054	0.047	0.056	0.053	0.054	0.054	0.058	0.057	0.056	0.059	0.06	0.058	0.055	0.056
Random	0.158	0.12	0.12	0.131	0.115	0.114	0.109	0.106	0.108	0.105	0.096	0.093	0.094	0.102	0.098	0.096	0.09	0.085	0.083	0.086
MDS - Certainty	0.0	0.043	0.039	0.035	0.032	0.039	0.037	0.05	0.046	0.047	0.044	0.042	0.047	0.045	0.044	0.043	0.046	0.045	0.047	0.052
MDS - Uncertainty	0.044	0.025	0.028	0.03	0.036	0.046	0.052	0.047	0.053	0.05	0.051	0.054	0.055	0.053	0.056	0.056	0.058	0.055	0.06	0.061
Hard Hybrid - Certainty	0.0	0.003	0.014	0.054	0.053	0.07	0.06	0.073	0.069	0.071	0.069	0.065	0.068	0.063	0.063	0.061	0.063	0.068	0.069	0.069
Hard Hybrid - Uncertainty	0.063	0.036	0.062	0.049	0.05	0.041	0.04	0.039	0.035	0.041	0.037	0.034	0.032	0.037	0.036	0.039	0.039	0.039	0.042	0.044
BCS	0.178	0.186	0.14	0.109	0.101	0.109	0.126	0.111	0.111	0.101	0.093	0.091	0.089	0.091	0.089	0.091	0.09	0.091	0.09	0.089

LCP_CaVousRegarde_2012-01-19_204700, Head Track: 371, Speaker Track: 290, OCR Name: 22 HvH 244X244 SvS 229X229 SvH 384



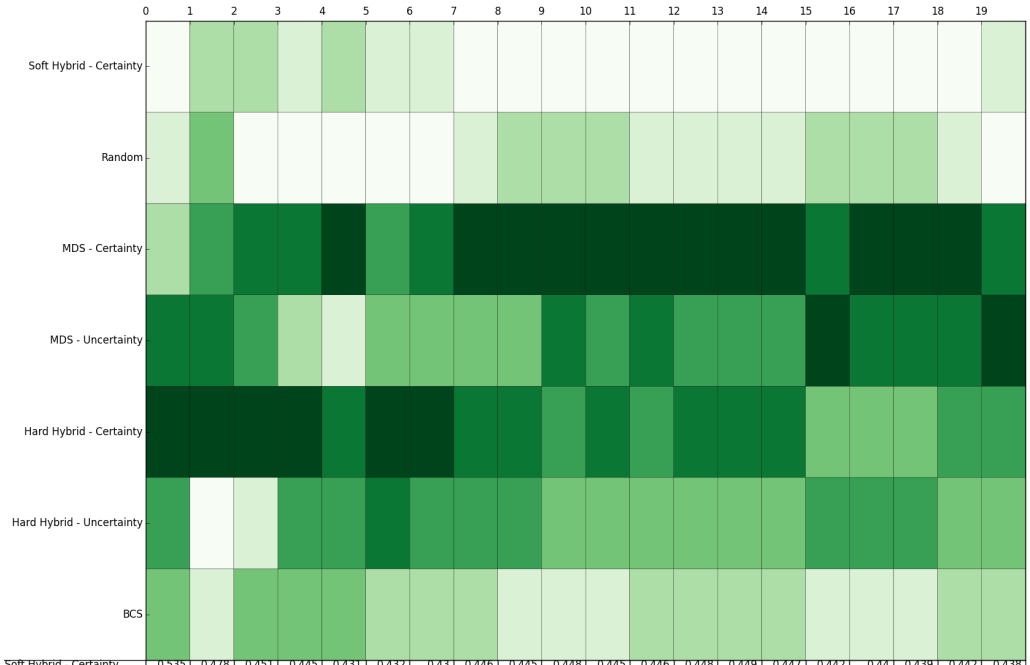
Soft Hybrid - Certainty	0.625	0.615	0.565	0.591	0.613	0.606	0.583	0.581	0.576	0.579	0.572	0.574	0.579	0.579	0.579	0.582	0.586	0.587	0.584	0.58
Random	0.462	0.477	0.488	0.52	0.508	0.518	0.526	0.524	0.524	0.527	0.533	0.534	0.523	0.527	0.535	0.544	0.538	0.538	0.539	0.535
MDS - Certainty	0.497	0.498	0.466	0.496	0.492	0.503	0.51	0.5	0.516	0.524	0.524	0.529	0.519	0.512	0.516	0.521	0.522	0.525	0.526	0.526
MDS - Uncertainty	0.462	0.488	0.495	0.534	0.537	0.556	0.554	0.545	0.538	0.531	0.529	0.528	0.53	0.525	0.522	0.525	0.529	0.525	0.525	0.525
Hard Hybrid - Certainty	0.569	0.552	0.566	0.567	0.564	0.57	0.567	0.568	0.567	0.566	0.573	0.577	0.572	0.572	0.571	0.574	0.574	0.576	0.574	0.571
Hard Hybrid - Uncertainty	0.501	0.511	0.572	0.59	0.598	0.599	0.61	0.616	0.607	0.602	0.599	0.602	0.597	0.591	0.593	0.593	0.591	0.586	0.586	0.585
BCS	0.685	0.643	0.597	0.598	0.582	0.577	0.572	0.579	0.583	0.581	0.579	0.572	0.563	0.56	0.554	0.549	0.555	0.555	0.552	0.552

LCP_CaVousRegarde_2012-01-25_204700, Head Track: 428, Speaker Track: 352, OCR Name: 25 HvH 281X281 SvS 282X282 SvH 429



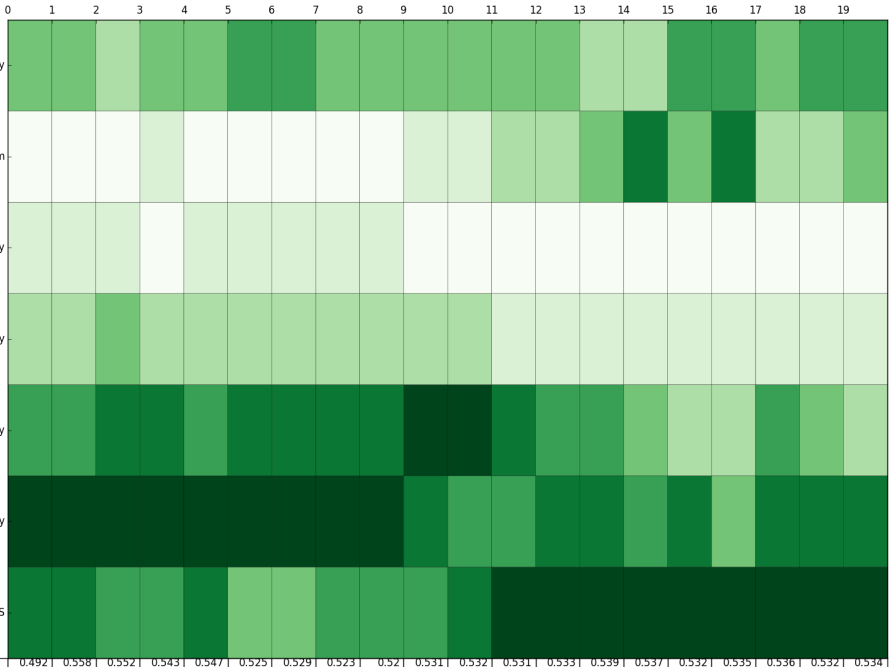
Soft Hybrid - Certainty	0.594	0.533	0.518	0.504	0.511	0.503	0.503	0.501	0.509	0.505	0.513	0.505	0.503	0.501	0.496	0.495	0.495	0.494	0.493	0.494
Random	0.575	0.54	0.505	0.493	0.485	0.496	0.497	0.498	0.499	0.503	0.497	0.502	0.509	0.501	0.499	0.503	0.502	0.498	0.496	0.495
MDS - Certainty	0.587	0.565	0.541	0.519	0.508	0.496	0.503	0.5	0.503	0.505	0.51	0.513	0.514	0.511	0.518	0.521	0.517	0.515	0.519	0.521
MDS - Uncertainty	0.489	0.469	0.445	0.463	0.462	0.481	0.503	0.506	0.508	0.505	0.508	0.505	0.503	0.504	0.506	0.505	0.503	0.503	0.507	0.511
Hard Hybrid - Certainty	0.495	0.492	0.496	0.513	0.521	0.517	0.513	0.519	0.527	0.527	0.522	0.521	0.514	0.516	0.517	0.52	0.517	0.515	0.519	0.52
Hard Hybrid - Uncertainty	0.558	0.558	0.567	0.554	0.546	0.538	0.533	0.523	0.514	0.507	0.511	0.516	0.516	0.514	0.511	0.508	0.506	0.513	0.515	0.515
BCS	0.588	0.532	0.512	0.506	0.516	0.525	0.526	0.526	0.533	0.533	0.53	0.529	0.531	0.528	0.527	0.527	0.528	0.524	0.523	0.522

LCP_EntreLesLignes_2011-12-16_192700, Head Track: 560, Speaker Track: 253, OCR Name: 15 HvH 345X345 SvS 203X203 SvH 493



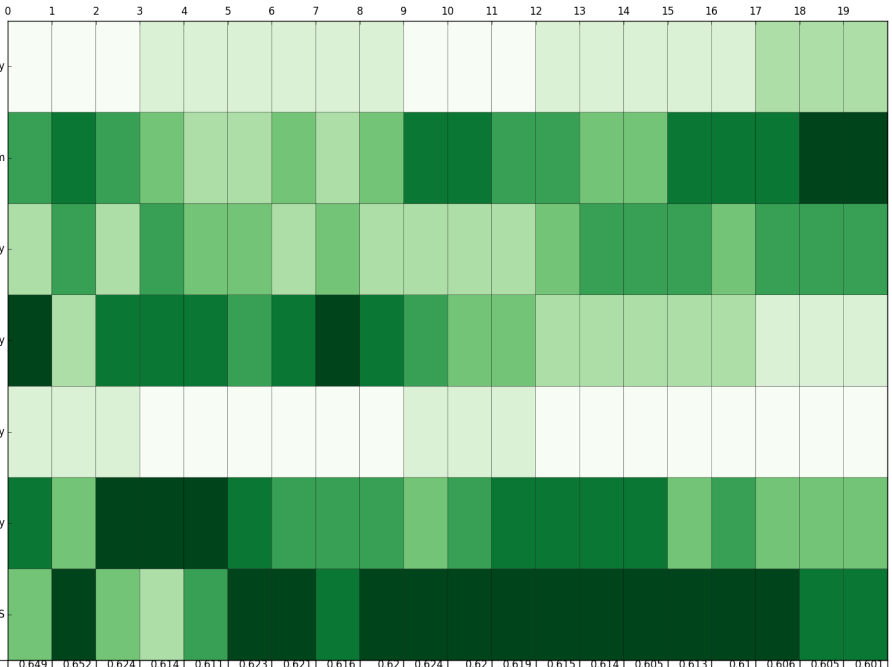
Soft Hybrid - Certainty	0.535	0.478	0.451	0.445	0.431	0.432	0.43	0.446	0.445	0.448	0.445	0.446	0.448	0.449	0.447	0.442	0.44	0.439	0.442	0.438
Random	0.48	0.457	0.464	0.459	0.438	0.437	0.434	0.432	0.435	0.425	0.425	0.437	0.445	0.441	0.442	0.433	0.435	0.434	0.437	0.439
MDS - Certainty	0.459	0.426	0.423	0.412	0.393	0.406	0.402	0.396	0.401	0.398	0.397	0.391	0.392	0.397	0.4	0.401	0.399	0.402	0.401	0.404
MDS - Uncertainty	0.422	0.424	0.436	0.438	0.434	0.424	0.418	0.422	0.413	0.405	0.409	0.402	0.406	0.407	0.404	0.397	0.4	0.402	0.402	0.402
Hard Hybrid - Certainty	0.336	0.417	0.395	0.403	0.399	0.397	0.395	0.401	0.401	0.407	0.406	0.408	0.401	0.402	0.403	0.404	0.41	0.406	0.408	0.412
Hard Hybrid - Uncertainty	0.451	0.494	0.455	0.418	0.411	0.401	0.409	0.402	0.404	0.412	0.415	0.422	0.415	0.416	0.415	0.404	0.408	0.406	0.41	0.413
BCS	0.452	0.492	0.45	0.421	0.428	0.428	0.423	0.425	0.436	0.435	0.438	0.434	0.428	0.429	0.43	0.437	0.439	0.437	0.435	0.435

LCP_EntreLesLignes_2012-01-27_192700, Head Track: 599, Speaker Track: 267, OCR Name: 14 HvH 329X329 SvS 210X210 SvH 460



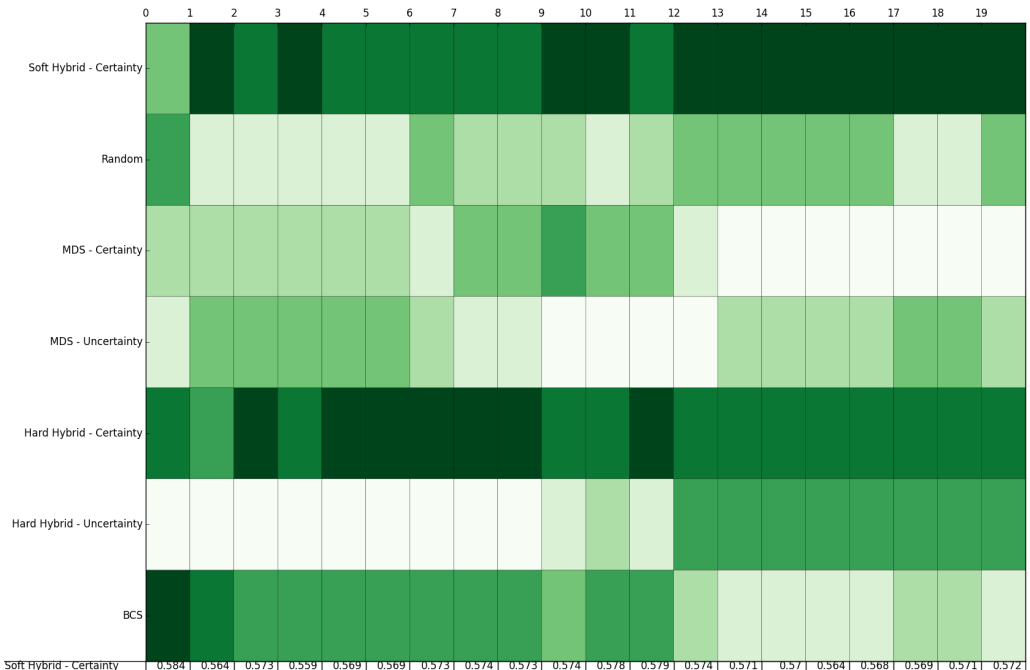
Soft Hybrid - Certainty	0.492	0.558	0.552	0.543	0.547	0.525	0.529	0.523	0.52	0.531	0.532	0.531	0.533	0.539	0.537	0.532	0.535	0.536	0.532	0.534
Random	0.664	0.642	0.597	0.564	0.584	0.561	0.561	0.573	0.562	0.553	0.547	0.54	0.539	0.533	0.529	0.533	0.535	0.538	0.538	0.536
MDS - Certainty	0.624	0.576	0.572	0.566	0.56	0.561	0.549	0.565	0.562	0.568	0.567	0.567	0.567	0.562	0.556	0.554	0.557	0.554	0.553	0.554
MDS - Uncertainty	0.544	0.565	0.525	0.551	0.547	0.551	0.538	0.524	0.538	0.54	0.542	0.547	0.544	0.544	0.545	0.547	0.549	0.55	0.551	0.552
Hard Hybrid - Certainty	0.484	0.508	0.508	0.509	0.522	0.517	0.516	0.512	0.507	0.514	0.509	0.523	0.528	0.529	0.536	0.537	0.539	0.536	0.537	0.536
Hard Hybrid - Uncertainty	0.416	0.45	0.432	0.455	0.492	0.49	0.486	0.498	0.504	0.515	0.526	0.524	0.525	0.526	0.53	0.531	0.539	0.534	0.531	0.531
BCS	0.44	0.49	0.511	0.523	0.52	0.537	0.536	0.518	0.52	0.518	0.515	0.522	0.525	0.522	0.517	0.518	0.527	0.524	0.527	0.528

LCP_EntreLesLignes_2012-05-11_192200, Head Track: 690, Speaker Track: 305, OCR Name: 13 HvH 310X310 SvS 240X240 SvH 596



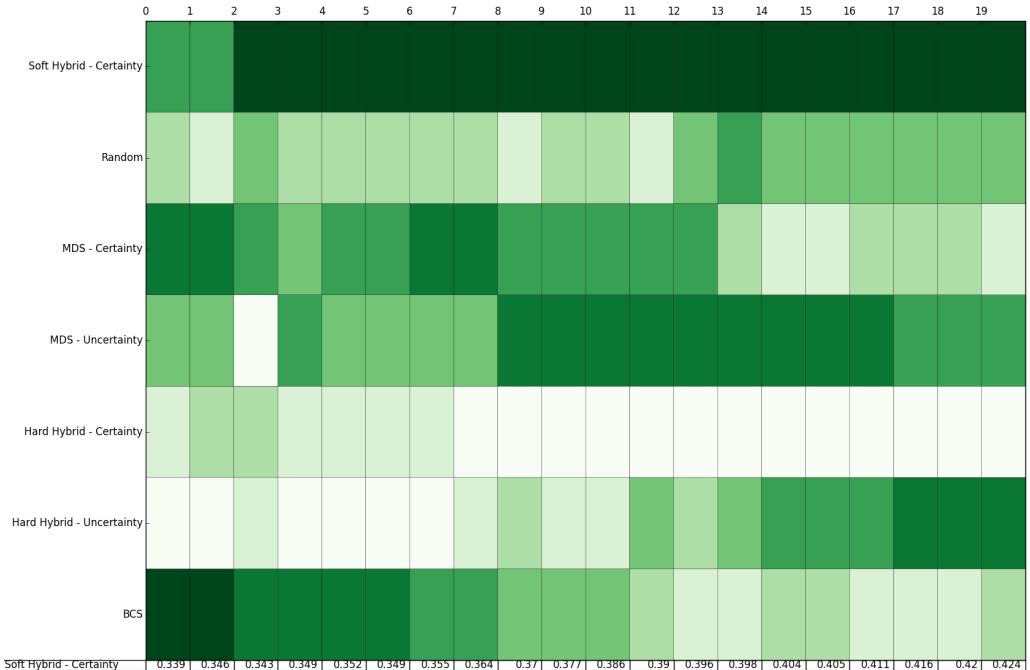
Soft Hybrid - Certainty	0.649	0.652	0.624	0.614	0.611	0.623	0.621	0.616	0.62	0.624	0.62	0.619	0.615	0.614	0.605	0.613	0.61	0.606	0.605	0.601
Random	0.578	0.555	0.577	0.586	0.584	0.594	0.586	0.598	0.594	0.582	0.585	0.59	0.587	0.591	0.59	0.586	0.587	0.59	0.589	0.587
MDS - Certainty	0.601	0.569	0.612	0.576	0.576	0.592	0.596	0.595	0.595	0.597	0.593	0.596	0.589	0.589	0.588	0.587	0.592	0.593	0.593	0.595
MDS - Uncertainty	0.543	0.57	0.571	0.562	0.568	0.578	0.581	0.575	0.583	0.584	0.588	0.591	0.594	0.6	0.6	0.601	0.602	0.608	0.607	0.603
Hard Hybrid - Certainty	0.608	0.593	0.612	0.624	0.632	0.644	0.637	0.62	0.623	0.616	0.614	0.611	0.616	0.618	0.625	0.621	0.622	0.618	0.617	0.615
Hard Hybrid - Uncertainty	0.567	0.569	0.555	0.527	0.542	0.574	0.584	0.586	0.589	0.594	0.586	0.586	0.585	0.582	0.583	0.59	0.588	0.595	0.595	0.596
BCS	0.582	0.541	0.578	0.594	0.575	0.569	0.576	0.58	0.578	0.581	0.578	0.576	0.58	0.578	0.578	0.577	0.583	0.587	0.591	0.594

LCP_LCPInfo13h30_2012-01-24_132700, Head Track: 212, Speaker Track: 163, OCR Name: 13 HvH 128X128 SvS 130X130 SvH 179



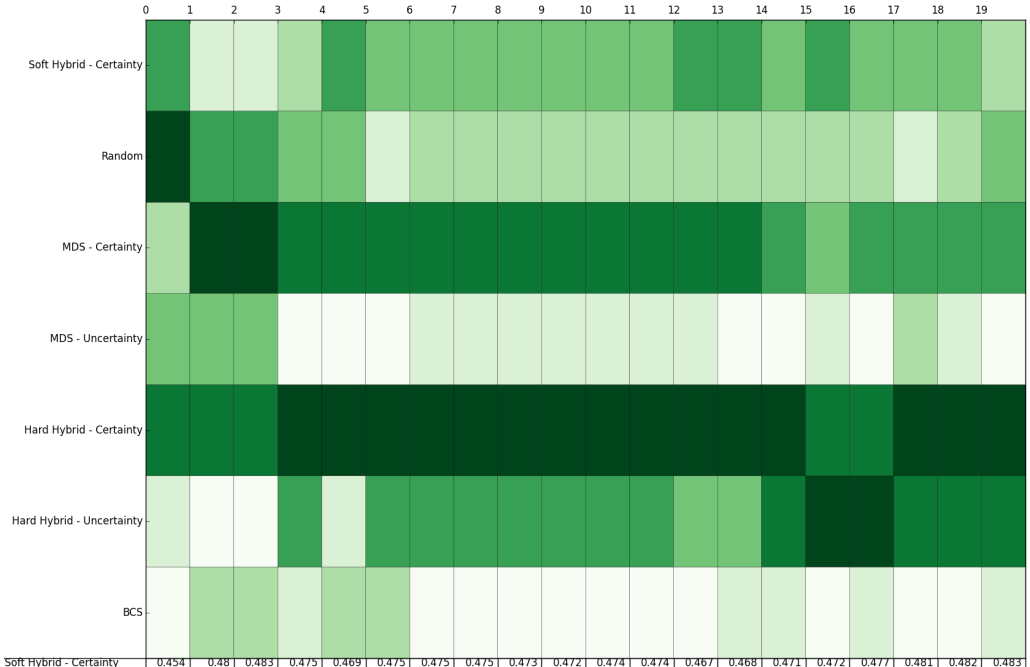
Soft Hybrid - Certainty	0.584	0.564	0.573	0.559	0.569	0.569	0.573	0.574	0.573	0.574	0.578	0.579	0.574	0.571	0.57	0.564	0.568	0.569	0.571	0.572
Random	0.572	0.614	0.628	0.615	0.612	0.604	0.604	0.605	0.6	0.604	0.602	0.597	0.596	0.596	0.599	0.596	0.597	0.599	0.602	0.601
MDS - Certainty	0.597	0.601	0.608	0.613	0.602	0.602	0.605	0.597	0.599	0.594	0.596	0.596	0.599	0.603	0.606	0.607	0.607	0.608	0.61	0.61
MDS - Uncertainty	0.611	0.598	0.596	0.594	0.592	0.598	0.604	0.605	0.602	0.604	0.604	0.606	0.603	0.6	0.599	0.596	0.597	0.598	0.599	0.601
Hard Hybrid - Certainty	0.569	0.575	0.557	0.561	0.559	0.562	0.564	0.573	0.572	0.58	0.581	0.579	0.581	0.578	0.579	0.582	0.583	0.582	0.583	0.586
Hard Hybrid - Uncertainty	0.652	0.646	0.632	0.633	0.631	0.628	0.611	0.61	0.613	0.604	0.6	0.6	0.593	0.595	0.592	0.592	0.592	0.595	0.592	0.592
BCS	0.525	0.565	0.59	0.582	0.585	0.588	0.593	0.594	0.591	0.598	0.595	0.595	0.598	0.601	0.603	0.601	0.6	0.599	0.601	0.603

LCP_LCPInfo13h30_2012-01-25_132700, Head Track: 331, Speaker Track: 172, OCR Name: 10 HvH 208X208 SvS 145X145 SvH 285



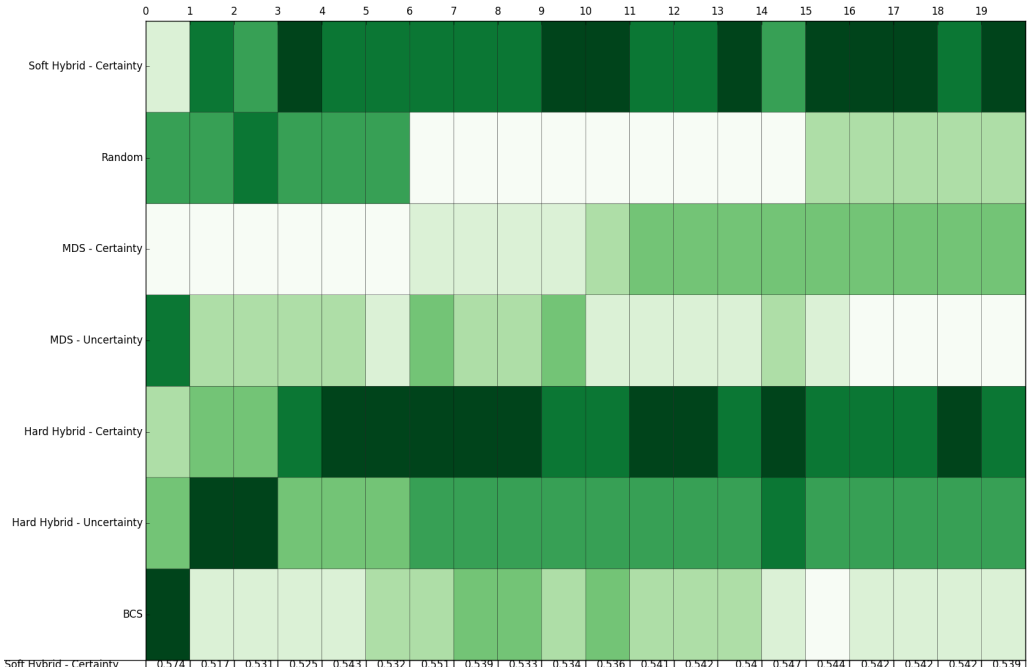
Soft Hybrid - Certainty	0.339	0.346	0.343	0.349	0.352	0.349	0.355	0.364	0.37	0.377	0.386	0.39	0.396	0.398	0.404	0.405	0.411	0.416	0.42	0.424
Random	0.351	0.36	0.36	0.368	0.379	0.385	0.393	0.398	0.403	0.403	0.405	0.411	0.413	0.414	0.418	0.419	0.426	0.43	0.429	0.432
MDS - Certainty	0.339	0.346	0.356	0.365	0.367	0.372	0.374	0.378	0.388	0.397	0.402	0.407	0.413	0.418	0.421	0.428	0.43	0.432	0.437	0.44
MDS - Uncertainty	0.339	0.354	0.368	0.361	0.368	0.374	0.385	0.39	0.387	0.391	0.393	0.396	0.397	0.404	0.409	0.415	0.42	0.424	0.428	0.431
Hard Hybrid - Certainty	0.367	0.358	0.366	0.371	0.379	0.389	0.393	0.401	0.404	0.411	0.412	0.418	0.422	0.425	0.43	0.435	0.437	0.441	0.446	0.451
Hard Hybrid - Uncertainty	0.374	0.379	0.368	0.382	0.388	0.392	0.396	0.4	0.402	0.404	0.405	0.408	0.413	0.414	0.415	0.418	0.422	0.424	0.427	0.43
BCS	0.334	0.34	0.353	0.36	0.366	0.372	0.379	0.388	0.392	0.399	0.405	0.41	0.415	0.419	0.421	0.428	0.432	0.437	0.438	0.44

LCP_LCPInfo13h30_2012-01-27_132700, Head Track: 306, Speaker Track: 177, OCR Name: 11 HvH 182X182 SvS 135X135 SvH 270



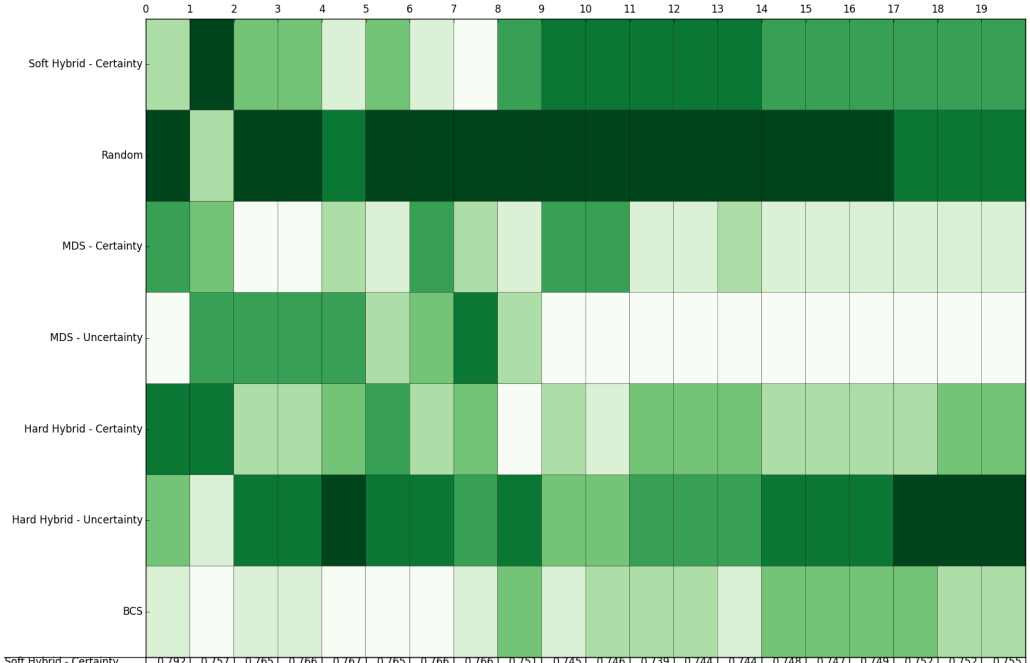
Soft Hybrid - Certainty	0.454	0.48	0.483	0.475	0.469	0.475	0.475	0.475	0.473	0.472	0.474	0.474	0.467	0.468	0.471	0.472	0.477	0.481	0.482	0.483
Random	0.358	0.468	0.462	0.465	0.479	0.492	0.488	0.478	0.484	0.479	0.485	0.481	0.484	0.488	0.485	0.486	0.486	0.488	0.483	0.483
MDS - Certainty	0.46	0.436	0.43	0.457	0.455	0.449	0.445	0.446	0.448	0.452	0.456	0.465	0.462	0.465	0.469	0.476	0.475	0.475	0.479	0.475
MDS - Uncertainty	0.459	0.473	0.482	0.499	0.505	0.504	0.492	0.491	0.485	0.491	0.489	0.481	0.487	0.492	0.49	0.489	0.493	0.487	0.491	0.494
Hard Hybrid - Certainty	0.434	0.455	0.454	0.438	0.436	0.436	0.428	0.435	0.429	0.434	0.447	0.448	0.451	0.454	0.458	0.459	0.462	0.46	0.459	0.458
Hard Hybrid - Uncertainty	0.515	0.534	0.485	0.464	0.484	0.474	0.468	0.468	0.468	0.467	0.465	0.465	0.474	0.47	0.463	0.459	0.46	0.465	0.466	0.47
BCS	0.544	0.478	0.483	0.479	0.483	0.489	0.503	0.504	0.506	0.509	0.5	0.495	0.491	0.492	0.489	0.49	0.491	0.489	0.492	0.491

LCP_PileEtFace_2011-11-19_085700, Head Track: 243, Speaker Track: 211, OCR Name: 40 HvH 173X173 SvS 162X162 SvH 289



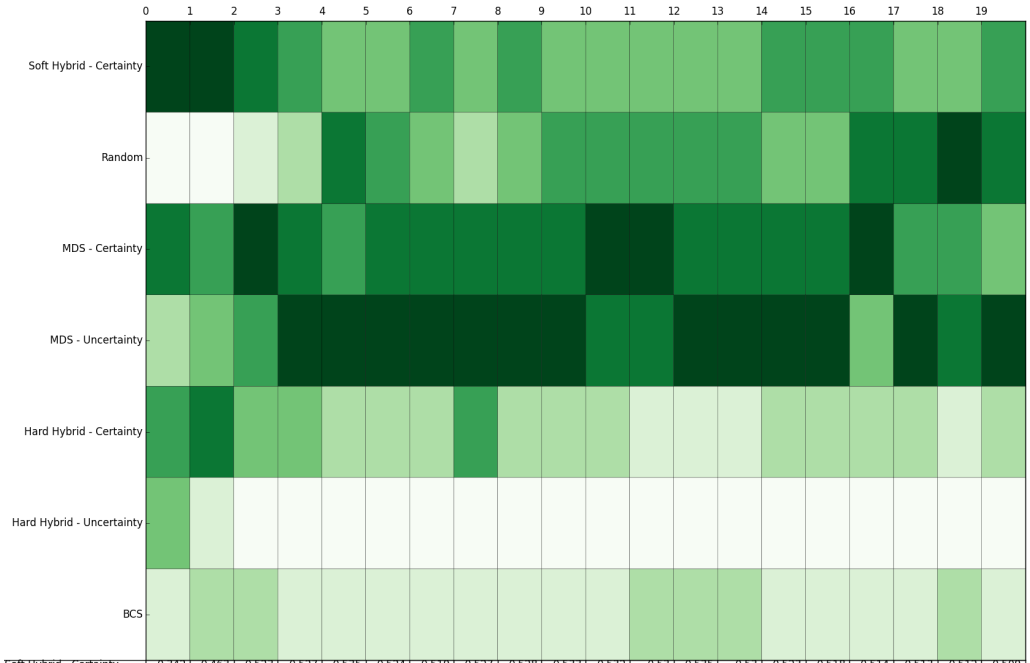
Soft Hybrid - Certainty	0.574	0.517	0.531	0.525	0.543	0.532	0.551	0.539	0.533	0.534	0.536	0.541	0.542	0.54	0.547	0.544	0.542	0.542	0.542	0.539
Random	0.547	0.541	0.524	0.547	0.56	0.561	0.582	0.595	0.587	0.588	0.583	0.586	0.579	0.575	0.578	0.57	0.568	0.568	0.565	0.57
MDS - Certainty	0.662	0.603	0.594	0.582	0.582	0.577	0.58	0.581	0.578	0.576	0.575	0.573	0.564	0.561	0.557	0.563	0.568	0.567	0.564	0.561
MDS - Uncertainty	0.504	0.543	0.558	0.56	0.571	0.576	0.572	0.576	0.576	0.563	0.575	0.578	0.575	0.573	0.571	0.57	0.576	0.575	0.58	0.575
Hard Hybrid - Certainty	0.557	0.542	0.554	0.537	0.524	0.516	0.517	0.524	0.527	0.535	0.54	0.541	0.535	0.544	0.546	0.55	0.548	0.544	0.54	0.542
Hard Hybrid - Uncertainty	0.55	0.506	0.518	0.552	0.571	0.564	0.556	0.545	0.555	0.562	0.567	0.56	0.551	0.548	0.547	0.552	0.554	0.553	0.545	0.548
BCS	0.483	0.588	0.579	0.577	0.573	0.572	0.576	0.57	0.566	0.565	0.568	0.574	0.575	0.572	0.576	0.572	0.572	0.574	0.573	0.571

LCP_PileEtFace_2011-12-01_192700, Head Track: 389, Speaker Track: 255, OCR Name: 51 HvH 250X250 SvS 197X197 SvH 349



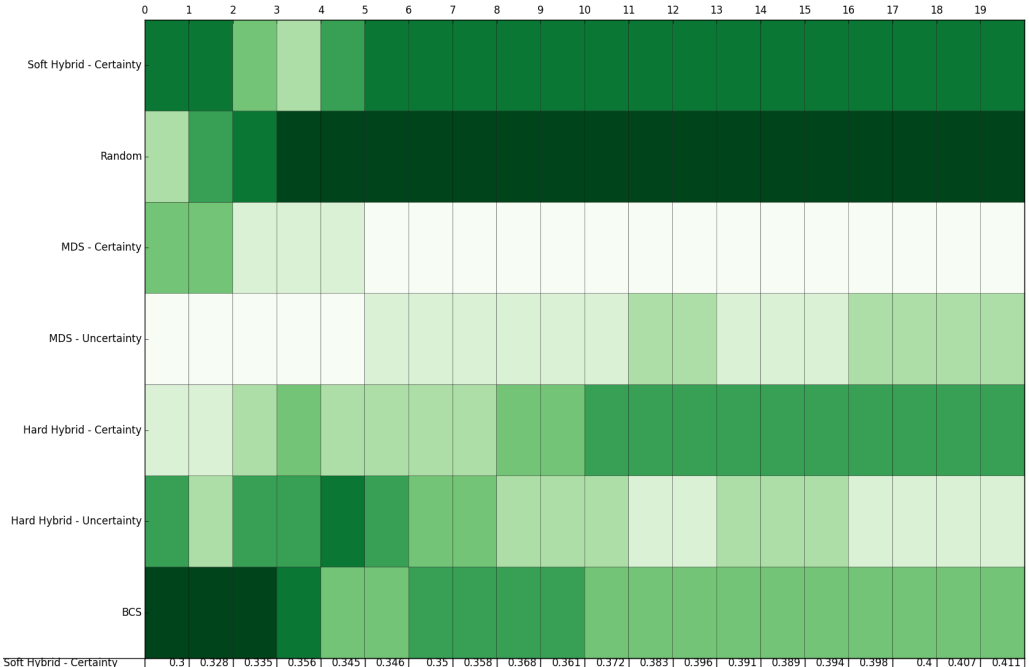
Soft Hybrid - Certainty	0.792	0.757	0.765	0.766	0.767	0.765	0.766	0.766	0.751	0.745	0.746	0.739	0.744	0.744	0.748	0.747	0.749	0.752	0.752	0.755
Random	0.742	0.777	0.729	0.728	0.743	0.733	0.723	0.721	0.728	0.726	0.732	0.738	0.743	0.739	0.741	0.745	0.745	0.747	0.748	0.751
MDS - Certainty	0.77	0.776	0.788	0.785	0.767	0.767	0.757	0.763	0.761	0.752	0.755	0.76	0.764	0.759	0.763	0.761	0.764	0.766	0.765	0.766
MDS - Uncertainty	0.804	0.774	0.756	0.76	0.754	0.765	0.759	0.756	0.76	0.763	0.764	0.767	0.767	0.771	0.773	0.772	0.774	0.773	0.774	0.775
Hard Hybrid - Certainty	0.757	0.769	0.782	0.775	0.766	0.763	0.762	0.759	0.761	0.757	0.761	0.757	0.757	0.757	0.76	0.759	0.761	0.762	0.758	0.758
Hard Hybrid - Uncertainty	0.771	0.778	0.753	0.732	0.738	0.743	0.748	0.758	0.751	0.754	0.756	0.756	0.752	0.748	0.744	0.745	0.747	0.745	0.742	0.744
BCS	0.796	0.792	0.785	0.779	0.775	0.775	0.767	0.765	0.755	0.761	0.759	0.759	0.763	0.764	0.757	0.757	0.757	0.757	0.759	0.76

LCP_PileEtFace_2012-01-12_192700, Head Track: 330, Speaker Track: 269, OCR Name: 52 HvH 228X228 SvS 218X218 SvH 373



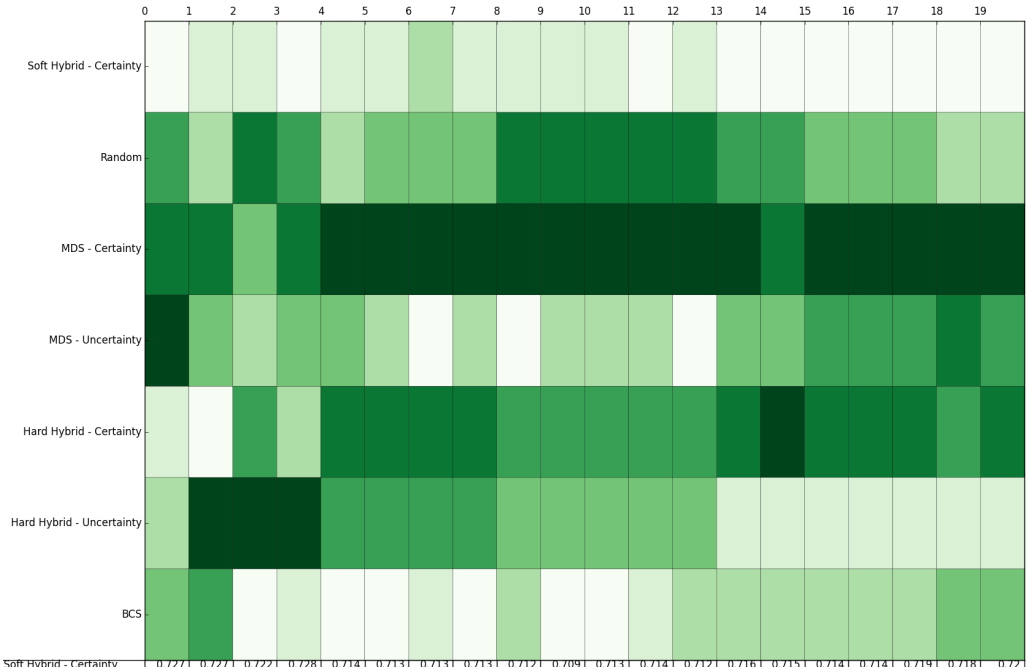
Soft Hybrid - Certainty	0.342	0.463	0.523	0.527	0.535	0.524	0.519	0.527	0.528	0.533	0.532	0.53	0.535	0.53	0.523	0.518	0.514	0.513	0.512	0.509
Random	0.643	0.591	0.593	0.556	0.517	0.523	0.533	0.533	0.531	0.517	0.524	0.521	0.524	0.527	0.533	0.523	0.513	0.505	0.5	0.504
MDS - Certainty	0.401	0.523	0.5	0.521	0.529	0.519	0.519	0.522	0.519	0.515	0.502	0.506	0.513	0.516	0.517	0.515	0.507	0.509	0.512	0.513
MDS - Uncertainty	0.573	0.542	0.537	0.509	0.498	0.501	0.514	0.512	0.505	0.506	0.514	0.517	0.509	0.507	0.5	0.507	0.514	0.504	0.502	0.503
Hard Hybrid - Certainty	0.441	0.496	0.548	0.532	0.536	0.555	0.536	0.522	0.54	0.542	0.546	0.557	0.553	0.549	0.545	0.541	0.541	0.539	0.543	0.54
Hard Hybrid - Uncertainty	0.555	0.589	0.621	0.615	0.614	0.609	0.611	0.604	0.591	0.589	0.591	0.58	0.577	0.572	0.556	0.556	0.549	0.552	0.553	0.563
BCS	0.62	0.58	0.573	0.586	0.601	0.564	0.561	0.573	0.575	0.571	0.559	0.547	0.537	0.545	0.547	0.552	0.544	0.541	0.542	0.546

LCP_PileEtFace_2012-01-19_192700, Head Track: 302, Speaker Track: 246, OCR Name: 43 HvH 222X222 SvS 191X191 SvH 290



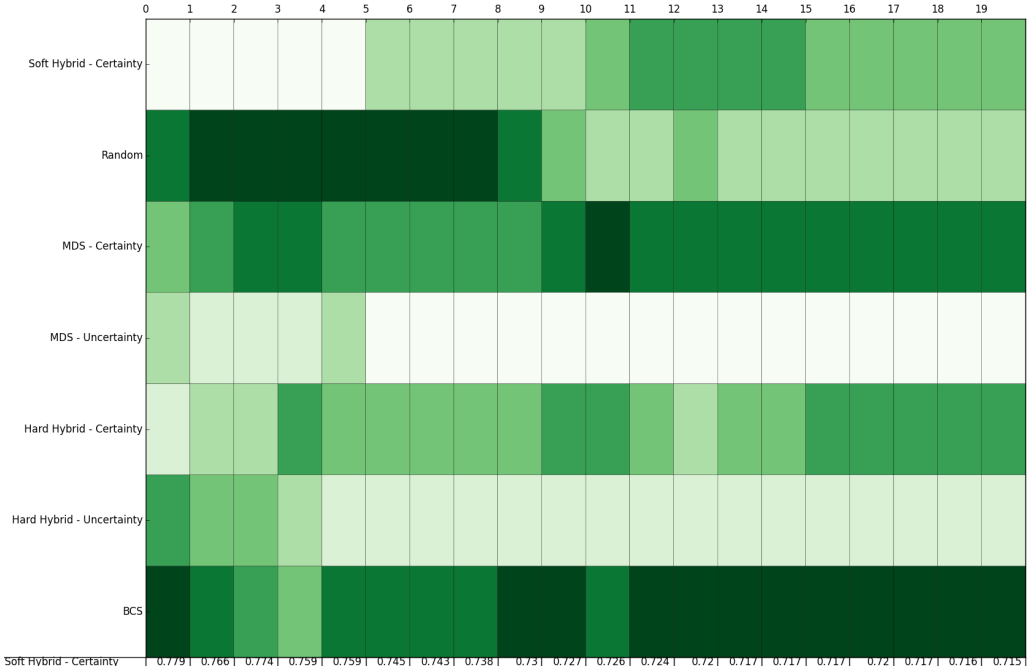
Soft Hybrid - Certainty	0.3	0.328	0.335	0.356	0.345	0.346	0.35	0.358	0.368	0.361	0.372	0.383	0.396	0.391	0.389	0.394	0.398	0.4	0.407	0.411
Random	0.393	0.34	0.324	0.328	0.326	0.312	0.323	0.33	0.344	0.351	0.348	0.347	0.361	0.366	0.375	0.378	0.377	0.383	0.386	0.389
MDS - Certainty	0.371	0.373	0.387	0.374	0.41	0.433	0.451	0.458	0.459	0.459	0.446	0.454	0.458	0.463	0.461	0.468	0.477	0.481	0.481	0.485
MDS - Uncertainty	0.467	0.514	0.463	0.439	0.413	0.41	0.423	0.43	0.418	0.419	0.424	0.416	0.421	0.434	0.438	0.442	0.442	0.452	0.458	0.461
Hard Hybrid - Certainty	0.42	0.396	0.367	0.353	0.373	0.385	0.395	0.399	0.4	0.405	0.405	0.401	0.4	0.398	0.398	0.405	0.411	0.412	0.418	0.427
Hard Hybrid - Uncertainty	0.348	0.39	0.327	0.349	0.343	0.362	0.381	0.385	0.408	0.416	0.42	0.424	0.423	0.429	0.435	0.44	0.454	0.459	0.464	0.47
BCS	0.248	0.252	0.308	0.342	0.367	0.368	0.38	0.374	0.385	0.404	0.411	0.411	0.406	0.409	0.414	0.419	0.42	0.424	0.428	0.435

LCP_PileEtFace_2012-01-26_192700, Head Track: 380, Speaker Track: 273, OCR Name: 50 HvH 223X223 SvS 207X207 SvH 349



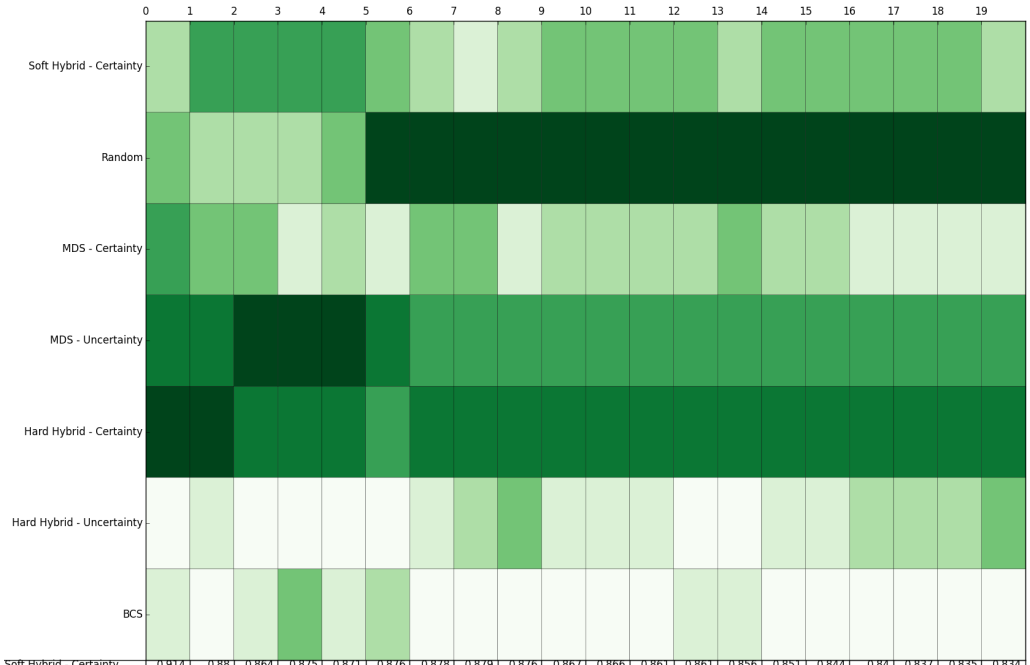
Soft Hybrid - Certainty	0.727	0.727	0.722	0.728	0.714	0.713	0.713	0.713	0.712	0.709	0.713	0.714	0.712	0.716	0.715	0.714	0.714	0.719	0.718	0.72
Random	0.643	0.708	0.685	0.696	0.706	0.7	0.7	0.704	0.695	0.689	0.692	0.687	0.689	0.693	0.697	0.701	0.7	0.705	0.707	0.712
MDS - Certainty	0.633	0.682	0.701	0.686	0.685	0.669	0.676	0.688	0.68	0.685	0.684	0.687	0.685	0.688	0.689	0.682	0.683	0.686	0.688	0.687
MDS - Uncertainty	0.581	0.705	0.719	0.696	0.697	0.707	0.717	0.707	0.722	0.708	0.705	0.708	0.713	0.706	0.702	0.699	0.695	0.694	0.692	0.693
Hard Hybrid - Certainty	0.723	0.741	0.69	0.703	0.685	0.677	0.686	0.694	0.699	0.693	0.694	0.694	0.696	0.69	0.688	0.693	0.692	0.691	0.693	0.687
Hard Hybrid - Uncertainty	0.711	0.656	0.658	0.679	0.697	0.689	0.696	0.703	0.701	0.696	0.698	0.698	0.705	0.71	0.713	0.707	0.707	0.708	0.712	0.715
BCS	0.672	0.694	0.733	0.72	0.718	0.714	0.714	0.716	0.705	0.717	0.714	0.711	0.71	0.707	0.713	0.707	0.703	0.707	0.707	0.708

LCP_TopQuestions_2012-01-25_232700, Head Track: 364, Speaker Track: 134, OCR Name: 11 HvH 204X204 SvS 102X102 SvH 310

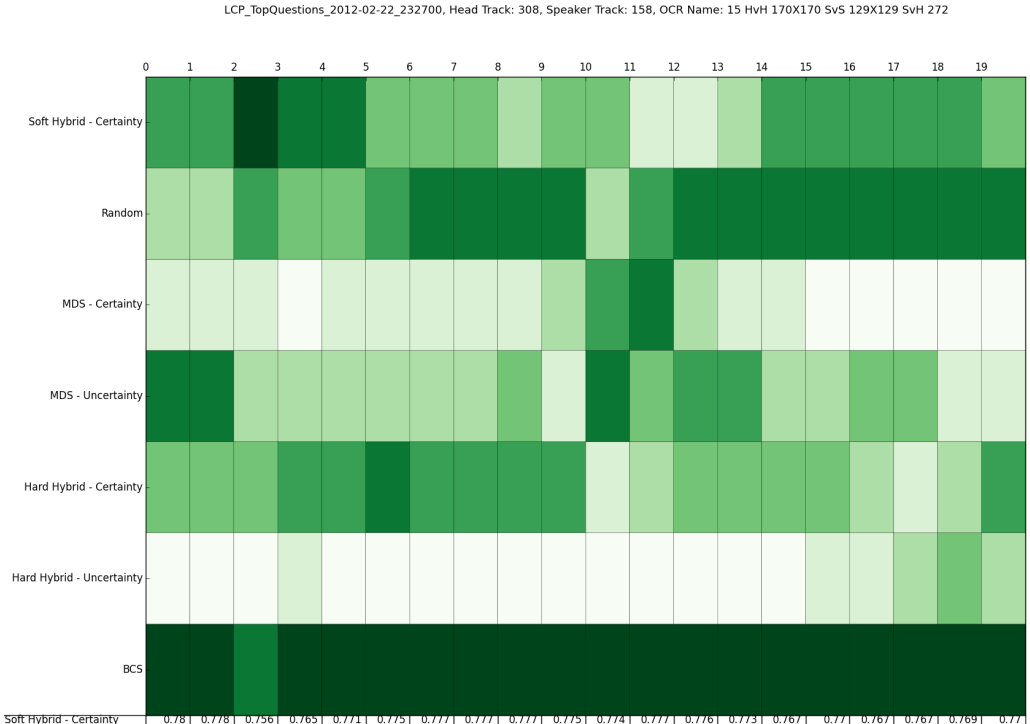


Soft Hybrid - Certainty	0.779	0.766	0.774	0.759	0.759	0.745	0.743	0.738	0.73	0.727	0.726	0.724	0.72	0.717	0.717	0.717	0.72	0.717	0.716	0.715
Random	0.712	0.673	0.68	0.694	0.695	0.708	0.709	0.709	0.717	0.724	0.728	0.728	0.726	0.728	0.728	0.731	0.73	0.732	0.732	0.731
MDS - Certainty	0.762	0.738	0.725	0.725	0.728	0.725	0.723	0.723	0.723	0.718	0.715	0.716	0.716	0.714	0.712	0.711	0.71	0.708	0.706	0.705
MDS - Uncertainty	0.766	0.763	0.741	0.748	0.75	0.759	0.755	0.749	0.752	0.75	0.749	0.746	0.745	0.743	0.744	0.743	0.742	0.738	0.738	0.736
Hard Hybrid - Certainty	0.767	0.746	0.74	0.734	0.74	0.742	0.733	0.73	0.728	0.72	0.724	0.726	0.729	0.725	0.719	0.717	0.715	0.711	0.708	0.709
Hard Hybrid - Uncertainty	0.757	0.74	0.739	0.736	0.752	0.75	0.75	0.744	0.746	0.743	0.739	0.734	0.734	0.734	0.737	0.737	0.734	0.736	0.733	0.733
BCS	0.709	0.715	0.726	0.735	0.719	0.72	0.718	0.718	0.712	0.715	0.716	0.713	0.712	0.708	0.706	0.707	0.709	0.706	0.704	0.705

LCP_TopQuestions_2012-02-14_232200, Head Track: 249, Speaker Track: 155, OCR Name: 13 HvH 133X133 SvS 127X127 SvH 272

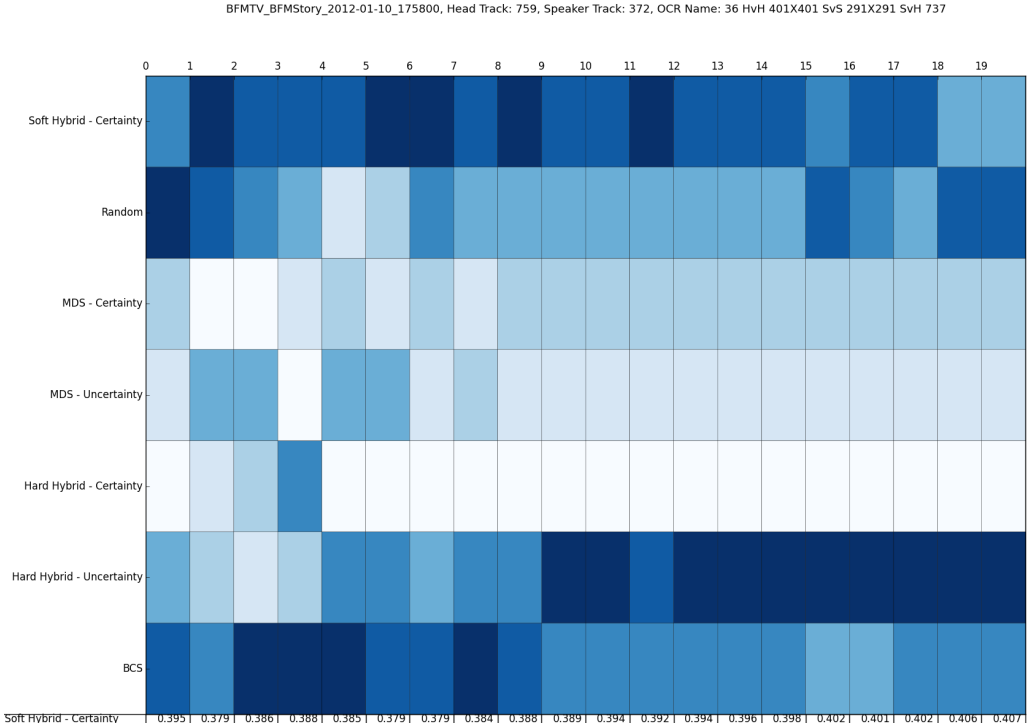


Soft Hybrid - Certainty	0.914	0.88	0.864	0.875	0.871	0.876	0.878	0.879	0.876	0.867	0.866	0.861	0.861	0.856	0.851	0.844	0.84	0.837	0.835	0.834
Random	0.897	0.896	0.896	0.892	0.875	0.849	0.842	0.834	0.836	0.839	0.842	0.841	0.841	0.833	0.834	0.83	0.829	0.827	0.826	0.824
MDS - Certainty	0.88	0.881	0.894	0.905	0.893	0.893	0.874	0.875	0.877	0.877	0.869	0.87	0.862	0.856	0.854	0.852	0.853	0.852	0.849	0.844
MDS - Uncertainty	0.859	0.855	0.861	0.864	0.86	0.865	0.87	0.871	0.861	0.855	0.853	0.847	0.848	0.844	0.84	0.841	0.839	0.834	0.834	0.832
Hard Hybrid - Certainty	0.858	0.848	0.864	0.871	0.87	0.868	0.863	0.861	0.854	0.851	0.851	0.841	0.843	0.84	0.838	0.835	0.834	0.83	0.829	0.827
Hard Hybrid - Uncertainty	0.962	0.905	0.914	0.916	0.908	0.897	0.889	0.878	0.872	0.877	0.873	0.87	0.873	0.868	0.861	0.855	0.847	0.844	0.837	0.834
BCS	0.929	0.937	0.913	0.891	0.895	0.893	0.89	0.894	0.894	0.891	0.883	0.877	0.873	0.867	0.863	0.861	0.857	0.857	0.851	0.847



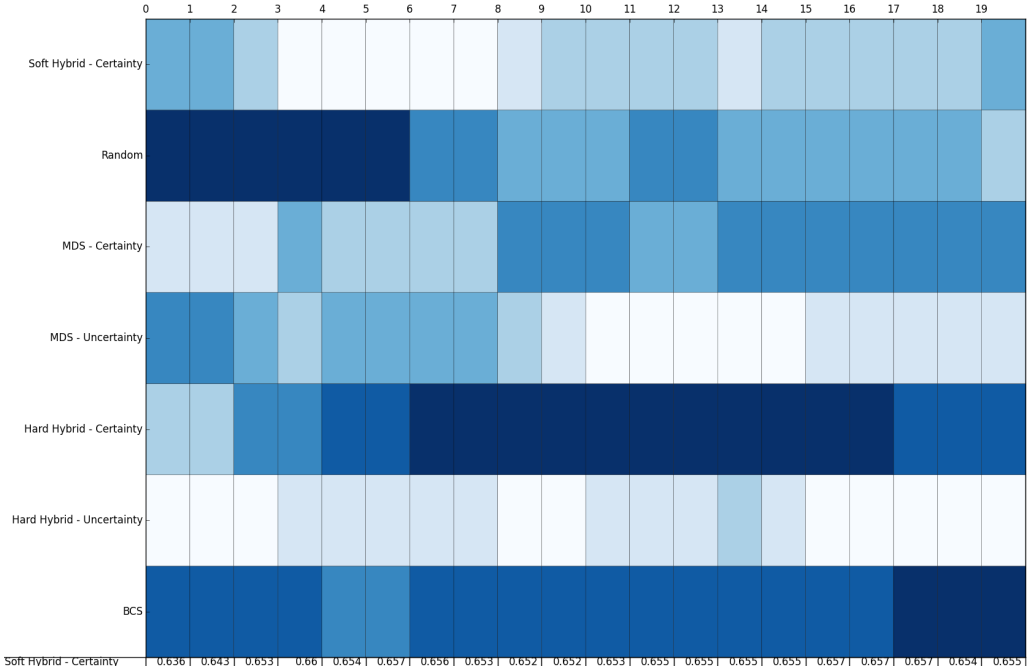
Soft Hybrid - Certainty	0.78	0.778	0.756	0.765	0.771	0.775	0.777	0.777	0.777	0.775	0.774	0.777	0.776	0.773	0.767	0.77	0.767	0.767	0.769	0.77
Random	0.784	0.798	0.785	0.783	0.785	0.771	0.766	0.769	0.772	0.773	0.775	0.773	0.768	0.766	0.764	0.766	0.767	0.767	0.766	0.769
MDS - Certainty	0.827	0.811	0.796	0.801	0.795	0.803	0.804	0.793	0.787	0.777	0.774	0.773	0.774	0.778	0.78	0.782	0.784	0.781	0.785	0.782
MDS - Uncertainty	0.762	0.777	0.79	0.787	0.785	0.785	0.781	0.782	0.776	0.779	0.773	0.773	0.772	0.769	0.774	0.771	0.771	0.772	0.773	0.774
Hard Hybrid - Certainty	0.782	0.791	0.788	0.779	0.779	0.77	0.766	0.77	0.775	0.775	0.776	0.776	0.773	0.772	0.769	0.771	0.774	0.774	0.773	0.769
Hard Hybrid - Uncertainty	0.845	0.818	0.809	0.79	0.802	0.804	0.806	0.811	0.805	0.8	0.796	0.794	0.789	0.787	0.783	0.78	0.779	0.774	0.773	0.772
BCS	0.714	0.762	0.76	0.752	0.759	0.761	0.76	0.752	0.751	0.757	0.756	0.756	0.755	0.758	0.753	0.759	0.76	0.76	0.762	0.763

Speaker score using face annotation is given blue heat maps.



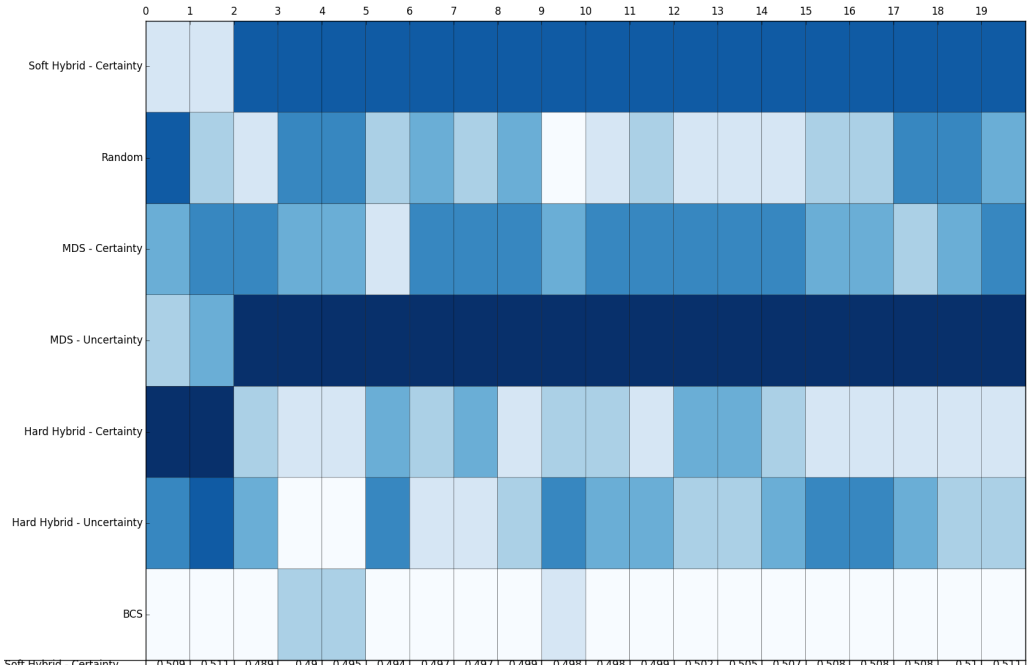
Soft Hybrid - Certainty	0.395	0.379	0.386	0.388	0.385	0.379	0.379	0.384	0.388	0.389	0.394	0.392	0.394	0.396	0.398	0.402	0.401	0.402	0.406	0.407
Random	0.36	0.385	0.396	0.403	0.404	0.402	0.399	0.402	0.403	0.405	0.406	0.403	0.404	0.404	0.403	0.4	0.402	0.403	0.404	0.405
MDS - Certainty	0.399	0.419	0.417	0.406	0.403	0.406	0.403	0.405	0.406	0.406	0.407	0.407	0.406	0.406	0.407	0.408	0.411	0.412	0.409	0.41
MDS - Uncertainty	0.405	0.391	0.402	0.407	0.399	0.401	0.406	0.404	0.411	0.411	0.409	0.411	0.411	0.415	0.416	0.418	0.416	0.418	0.418	0.417
Hard Hybrid - Certainty	0.406	0.412	0.413	0.402	0.406	0.408	0.41	0.416	0.418	0.423	0.422	0.421	0.42	0.421	0.419	0.423	0.426	0.427	0.425	0.424
Hard Hybrid - Uncertainty	0.397	0.412	0.415	0.405	0.39	0.394	0.401	0.398	0.399	0.389	0.39	0.393	0.391	0.392	0.394	0.394	0.396	0.396	0.396	0.397
BCS	0.392	0.386	0.36	0.369	0.379	0.383	0.384	0.384	0.39	0.393	0.395	0.4	0.402	0.401	0.4	0.405	0.404	0.403	0.406	0.407

BFMTV_BFMStory_2012-01-23_175800, Head Track: 782, Speaker Track: 338, OCR Name: 34 HvH 406X406 SvS 247X247 SvH 556



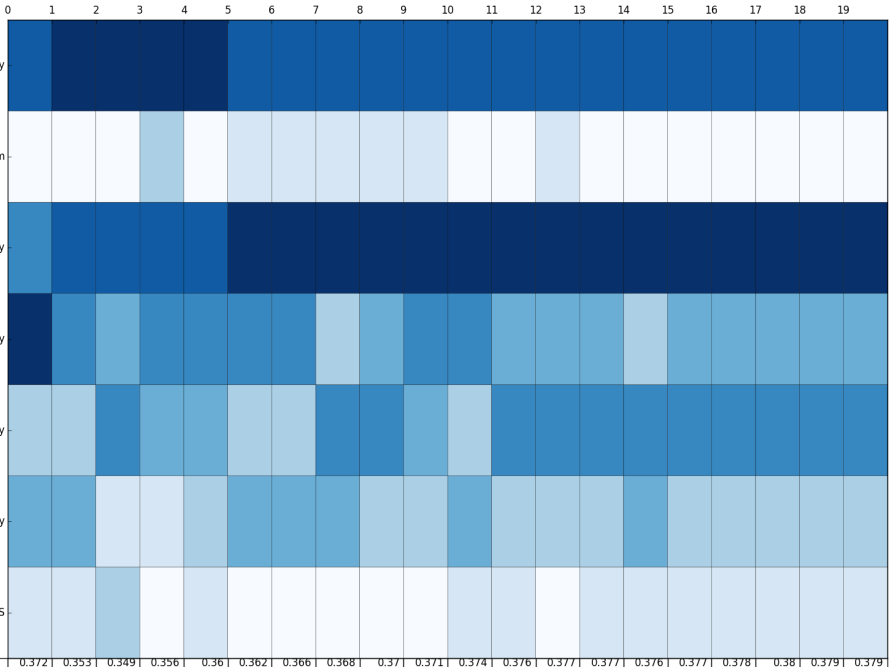
Soft Hybrid - Certainty	0.636	0.643	0.653	0.66	0.654	0.657	0.656	0.653	0.652	0.652	0.653	0.655	0.655	0.655	0.655	0.657	0.657	0.657	0.654	0.655
Random	0.587	0.595	0.617	0.623	0.631	0.634	0.639	0.642	0.644	0.647	0.646	0.645	0.647	0.65	0.651	0.651	0.651	0.653	0.654	0.656
MDS - Certainty	0.659	0.658	0.659	0.651	0.647	0.648	0.65	0.648	0.644	0.645	0.644	0.646	0.65	0.65	0.651	0.65	0.649	0.65	0.649	0.65
MDS - Uncertainty	0.63	0.63	0.647	0.653	0.637	0.646	0.646	0.647	0.65	0.654	0.66	0.658	0.66	0.658	0.658	0.657	0.658	0.66	0.661	0.662
Hard Hybrid - Certainty	0.638	0.646	0.631	0.635	0.631	0.634	0.635	0.634	0.632	0.632	0.633	0.634	0.635	0.637	0.639	0.64	0.641	0.643	0.644	0.647
Hard Hybrid - Uncertainty	0.683	0.662	0.667	0.655	0.65	0.652	0.65	0.652	0.653	0.655	0.655	0.658	0.657	0.654	0.658	0.661	0.661	0.663	0.666	0.667
BCS	0.611	0.613	0.629	0.628	0.633	0.636	0.639	0.636	0.637	0.639	0.638	0.64	0.637	0.64	0.641	0.642	0.643	0.643	0.643	0.642

BFMTV_BFMStory_2012-02-14_175800, Head Track: 753, Speaker Track: 348, OCR Name: 33 HvH 405X405 SvS 272X272 SvH 643



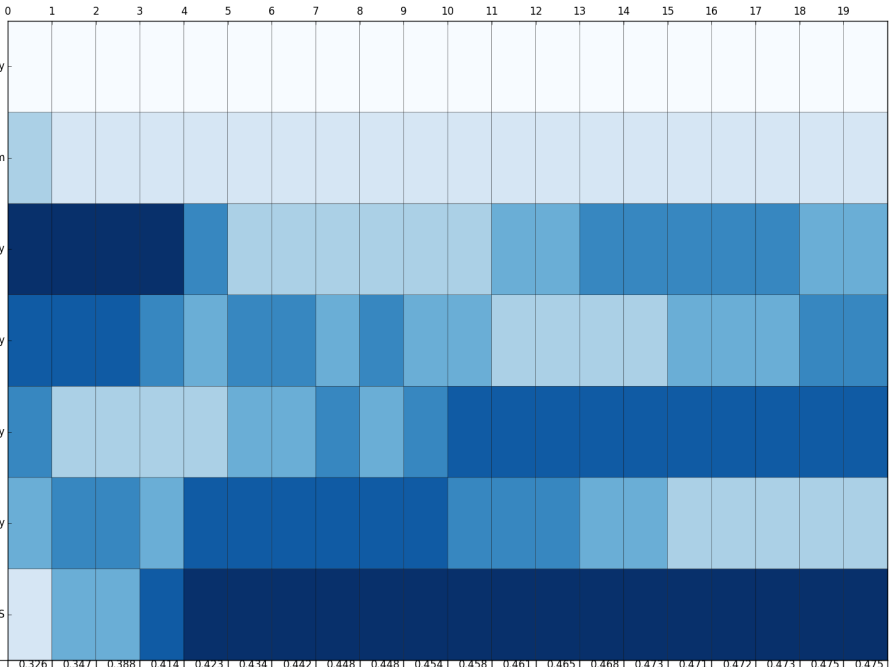
Soft Hybrid - Certainty	0.509	0.511	0.489	0.49	0.495	0.494	0.497	0.497	0.499	0.498	0.498	0.499	0.502	0.505	0.507	0.508	0.508	0.508	0.51	0.511
Random	0.48	0.51	0.508	0.497	0.5	0.506	0.507	0.51	0.508	0.512	0.514	0.515	0.517	0.517	0.518	0.516	0.516	0.515	0.516	0.517
MDS - Certainty	0.499	0.495	0.501	0.499	0.504	0.508	0.503	0.501	0.505	0.509	0.51	0.513	0.512	0.512	0.513	0.515	0.515	0.517	0.517	0.517
MDS - Uncertainty	0.508	0.504	0.472	0.476	0.476	0.468	0.477	0.476	0.474	0.475	0.479	0.477	0.48	0.483	0.488	0.487	0.489	0.491	0.494	0.495
Hard Hybrid - Certainty	0.474	0.489	0.502	0.509	0.509	0.505	0.508	0.509	0.511	0.509	0.512	0.516	0.513	0.514	0.516	0.518	0.518	0.52	0.521	0.523
Hard Hybrid - Uncertainty	0.484	0.49	0.501	0.518	0.518	0.505	0.509	0.511	0.51	0.508	0.511	0.513	0.514	0.515	0.515	0.511	0.511	0.516	0.519	0.519
BCS	0.54	0.526	0.514	0.506	0.509	0.509	0.512	0.518	0.514	0.512	0.516	0.519	0.519	0.522	0.524	0.526	0.526	0.526	0.527	0.526

BFMTV_BFMStory_2012-02-20_175800, Head Track: 871, Speaker Track: 342, OCR Name: 30 HvH 450X450 SvS 267X267 SvH 710



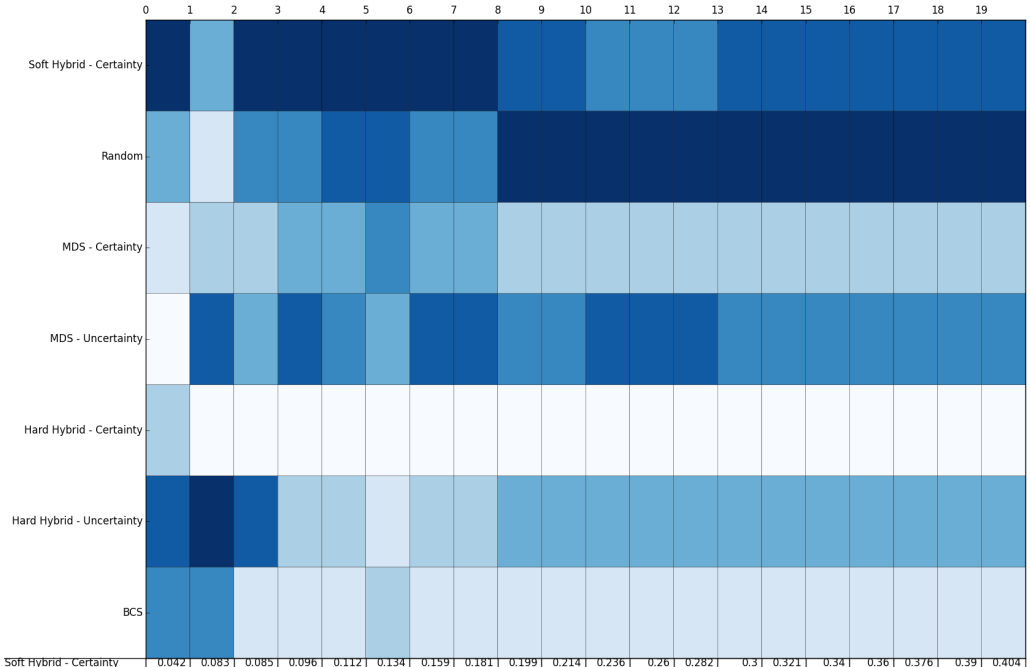
Soft Hybrid - Certainty	0.372	0.353	0.349	0.356	0.36	0.362	0.366	0.368	0.37	0.371	0.374	0.376	0.377	0.377	0.376	0.377	0.378	0.38	0.379	0.379
Random	0.407	0.393	0.391	0.385	0.393	0.39	0.392	0.393	0.393	0.394	0.397	0.399	0.398	0.4	0.402	0.401	0.404	0.407	0.407	0.406
MDS - Certainty	0.373	0.358	0.365	0.362	0.364	0.362	0.364	0.362	0.361	0.363	0.366	0.367	0.368	0.368	0.369	0.37	0.372	0.374	0.376	0.376
MDS - Uncertainty	0.369	0.369	0.372	0.376	0.373	0.372	0.374	0.379	0.377	0.375	0.376	0.38	0.381	0.383	0.385	0.385	0.388	0.388	0.387	0.388
Hard Hybrid - Certainty	0.387	0.381	0.371	0.377	0.376	0.38	0.378	0.376	0.376	0.379	0.38	0.378	0.379	0.38	0.383	0.385	0.387	0.387	0.386	0.387
Hard Hybrid - Uncertainty	0.376	0.373	0.385	0.386	0.383	0.38	0.378	0.377	0.382	0.38	0.38	0.382	0.385	0.384	0.385	0.392	0.393	0.39	0.39	0.389
BCS	0.388	0.383	0.377	0.387	0.385	0.391	0.394	0.394	0.395	0.396	0.396	0.398	0.399	0.4	0.4	0.4	0.401	0.399	0.399	0.398

BFMTV_CultureEtVous_2012-01-13_065040, Head Track: 69, Speaker Track: 38, OCR Name: 3 HvH 29X29 SvS 28X28 SvH 42



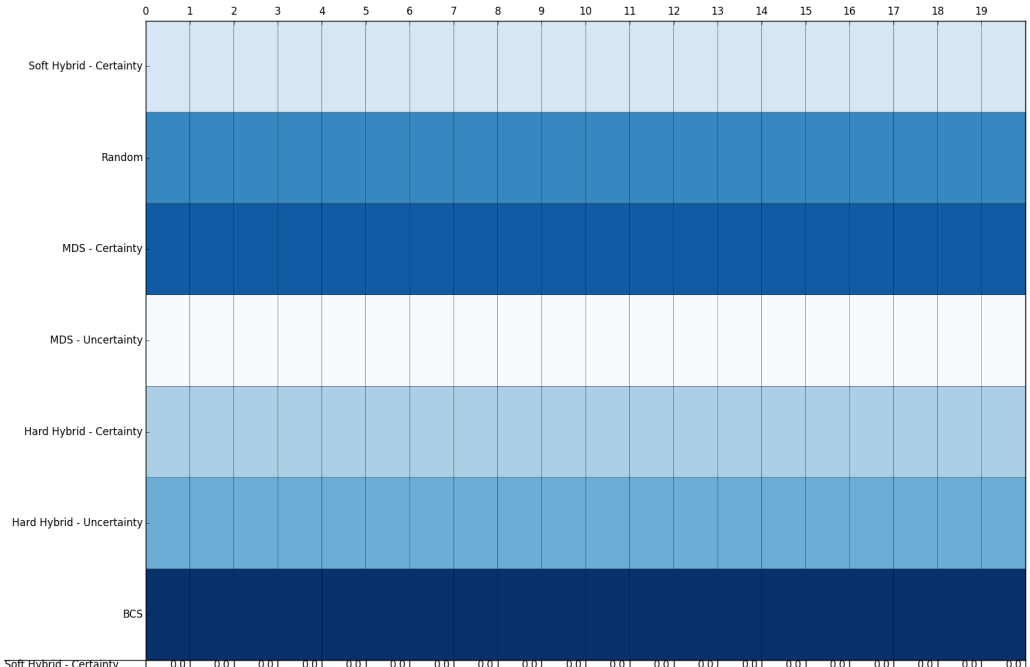
Soft Hybrid - Certainty	0.326	0.347	0.388	0.414	0.423	0.434	0.442	0.448	0.448	0.454	0.458	0.461	0.465	0.468	0.473	0.471	0.472	0.473	0.475	0.475
Random	0.324	0.339	0.334	0.338	0.346	0.356	0.363	0.371	0.377	0.385	0.394	0.401	0.408	0.413	0.418	0.42	0.424	0.427	0.43	0.43
MDS - Certainty	0.293	0.293	0.293	0.292	0.302	0.316	0.318	0.33	0.335	0.34	0.346	0.354	0.362	0.374	0.384	0.392	0.401	0.407	0.411	0.417
MDS - Uncertainty	0.293	0.293	0.293	0.299	0.302	0.299	0.302	0.309	0.312	0.327	0.344	0.359	0.366	0.377	0.385	0.393	0.402	0.407	0.41	0.414
Hard Hybrid - Certainty	0.322	0.324	0.324	0.324	0.318	0.314	0.307	0.308	0.315	0.314	0.328	0.344	0.353	0.363	0.37	0.379	0.388	0.395	0.4	0.406
Hard Hybrid - Uncertainty	0.323	0.323	0.312	0.306	0.295	0.288	0.283	0.282	0.293	0.314	0.331	0.345	0.36	0.374	0.385	0.396	0.406	0.415	0.422	0.425
BCS	0.326	0.324	0.314	0.294	0.288	0.284	0.281	0.279	0.281	0.289	0.301	0.316	0.333	0.348	0.356	0.366	0.374	0.383	0.39	0.396

BFMTV_CultureEtVous_2012-01-16_075000, Head Track: 163, Speaker Track: 42, OCR Name: 5 HvH 78X78 SvS 32X32 SvH 117



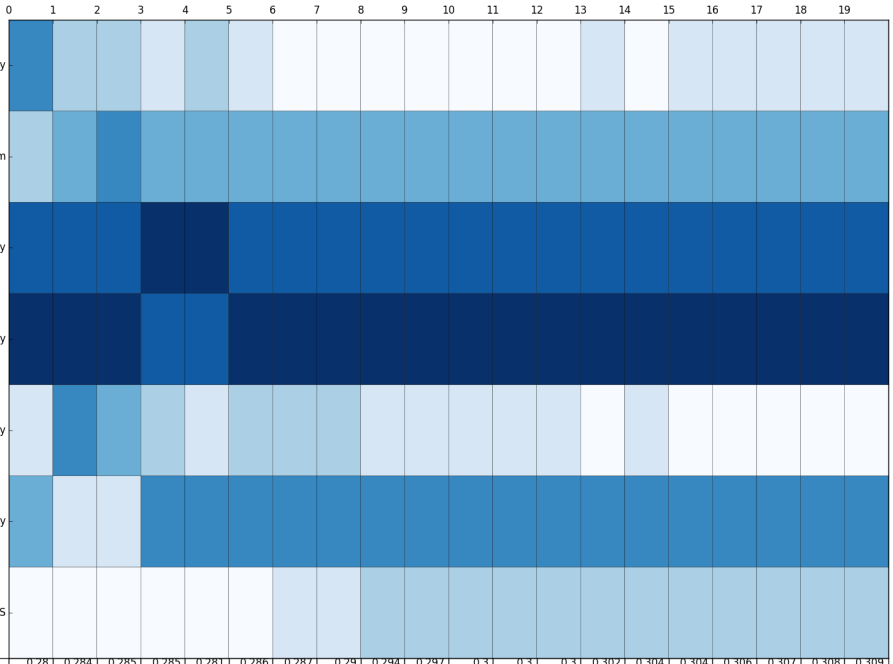
Soft Hybrid - Certainty	0.042	0.083	0.085	0.096	0.112	0.134	0.159	0.181	0.199	0.214	0.236	0.26	0.282	0.3	0.321	0.34	0.36	0.376	0.39	0.404
Random	0.086	0.113	0.104	0.117	0.127	0.145	0.175	0.188	0.198	0.208	0.221	0.234	0.256	0.283	0.295	0.312	0.329	0.351	0.368	0.384
MDS - Certainty	0.092	0.108	0.12	0.126	0.148	0.151	0.186	0.204	0.243	0.261	0.286	0.314	0.339	0.361	0.376	0.39	0.401	0.414	0.431	0.441
MDS - Uncertainty	0.092	0.078	0.107	0.115	0.139	0.154	0.167	0.182	0.206	0.223	0.231	0.252	0.275	0.302	0.328	0.348	0.364	0.376	0.391	0.404
Hard Hybrid - Certainty	0.089	0.14	0.172	0.206	0.239	0.292	0.307	0.344	0.364	0.381	0.401	0.418	0.435	0.448	0.457	0.468	0.481	0.49	0.499	0.511
Hard Hybrid - Uncertainty	0.049	0.078	0.095	0.126	0.173	0.196	0.209	0.216	0.232	0.242	0.256	0.284	0.309	0.328	0.345	0.356	0.377	0.399	0.414	0.427
BCS	0.076	0.083	0.134	0.159	0.177	0.187	0.219	0.246	0.269	0.305	0.338	0.365	0.394	0.419	0.441	0.455	0.472	0.482	0.493	0.499

BFMTV_CultureEtVous_2012-01-17_065040, Head Track: 48, Speaker Track: 49, OCR Name: 5 HvH 23X23 SvS 35X35 SvH 30



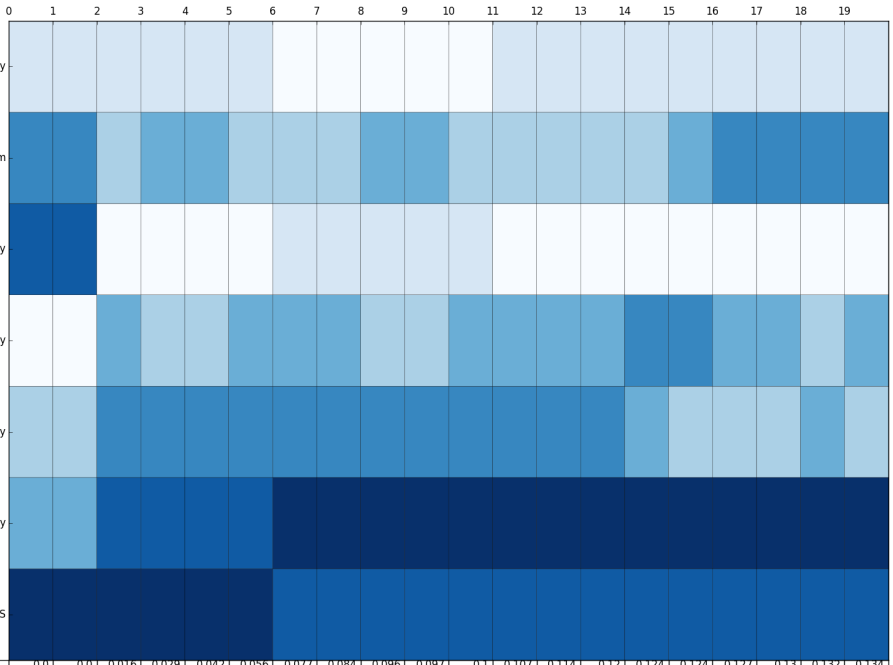
Soft Hybrid - Certainty	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Random	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
MDS - Certainty	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
MDS - Uncertainty	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Hard Hybrid - Certainty	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Hard Hybrid - Uncertainty	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
BCS	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

BFMTV_CultureEtVous_2012-01-18_065040, Head Track: 102, Speaker Track: 47, OCR Name: 6 HvH 41X41 SvS 37X37 SvH 42



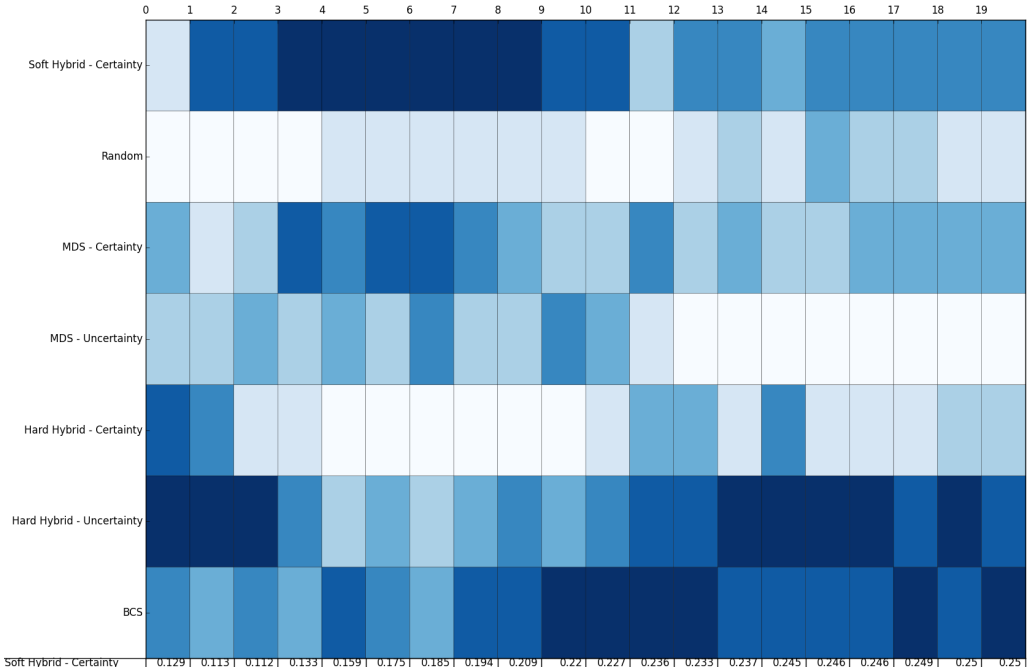
Soft Hybrid - Certainty	0.28	0.284	0.285	0.285	0.281	0.286	0.287	0.29	0.294	0.297	0.3	0.3	0.3	0.302	0.304	0.304	0.306	0.307	0.308	0.309
Random	0.285	0.279	0.277	0.275	0.278	0.275	0.277	0.279	0.281	0.284	0.287	0.29	0.291	0.292	0.295	0.295	0.296	0.297	0.296	0.297
MDS - Certainty	0.258	0.268	0.271	0.26	0.264	0.27	0.268	0.274	0.274	0.274	0.277	0.28	0.283	0.283	0.284	0.287	0.289	0.29	0.291	0.292
MDS - Uncertainty	0.253	0.267	0.268	0.261	0.266	0.267	0.267	0.263	0.262	0.266	0.267	0.269	0.273	0.278	0.278	0.28	0.282	0.284	0.286	0.288
Hard Hybrid - Certainty	0.288	0.275	0.279	0.284	0.283	0.282	0.279	0.284	0.288	0.292	0.295	0.297	0.3	0.303	0.304	0.306	0.308	0.309	0.311	0.31
Hard Hybrid - Uncertainty	0.284	0.295	0.286	0.273	0.276	0.274	0.276	0.275	0.275	0.277	0.282	0.285	0.289	0.29	0.291	0.289	0.292	0.294	0.295	0.295
BCS	0.296	0.305	0.301	0.296	0.288	0.288	0.284	0.286	0.285	0.289	0.292	0.293	0.294	0.296	0.298	0.301	0.301	0.303	0.304	0.305

BFMTV_CultureEtVous_2012-01-19_065040, Head Track: 78, Speaker Track: 41, OCR Name: 4 HvH 43X43 SvS 31X31 SvH 52



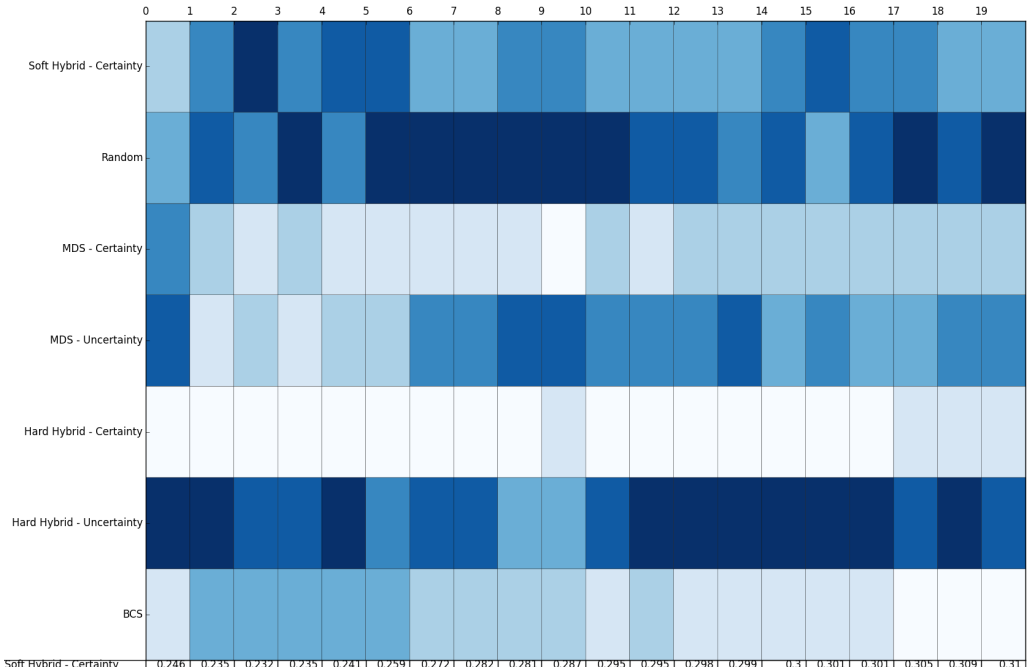
Soft Hybrid - Certainty	0.0	0.0	0.016	0.029	0.042	0.056	0.077	0.084	0.096	0.097	0.1	0.107	0.114	0.12	0.124	0.124	0.127	0.13	0.132	0.134
Random	0.0	0.0	0.007	0.009	0.016	0.041	0.054	0.064	0.066	0.074	0.083	0.088	0.094	0.098	0.102	0.109	0.109	0.114	0.114	0.118
MDS - Certainty	0.0	0.0	0.031	0.041	0.048	0.061	0.07	0.079	0.079	0.085	0.098	0.112	0.119	0.127	0.129	0.135	0.135	0.137	0.138	0.142
MDS - Uncertainty	0.0	0.0	0.006	0.027	0.032	0.036	0.047	0.056	0.067	0.075	0.082	0.085	0.091	0.094	0.098	0.106	0.111	0.117	0.124	0.127
Hard Hybrid - Certainty	0.0	0.0	0.0	0.0	0.004	0.006	0.017	0.028	0.041	0.042	0.055	0.07	0.084	0.092	0.1	0.109	0.115	0.12	0.124	0.129
Hard Hybrid - Uncertainty	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.006	0.013	0.016	0.026	0.039	0.051	0.062	0.07	0.08	0.087	0.094	0.097	0.101
BCS	0.0	0.0	0.0	0.0	0.0	0.0	0.004	0.015	0.025	0.035	0.043	0.056	0.065	0.077	0.081	0.088	0.094	0.1	0.103	0.106

BFMTV_CultureEtVous_2012-02-14_065040, Head Track: 104, Speaker Track: 48, OCR Name: 9 HvH 48X48 SvS 39X39 SvH 61



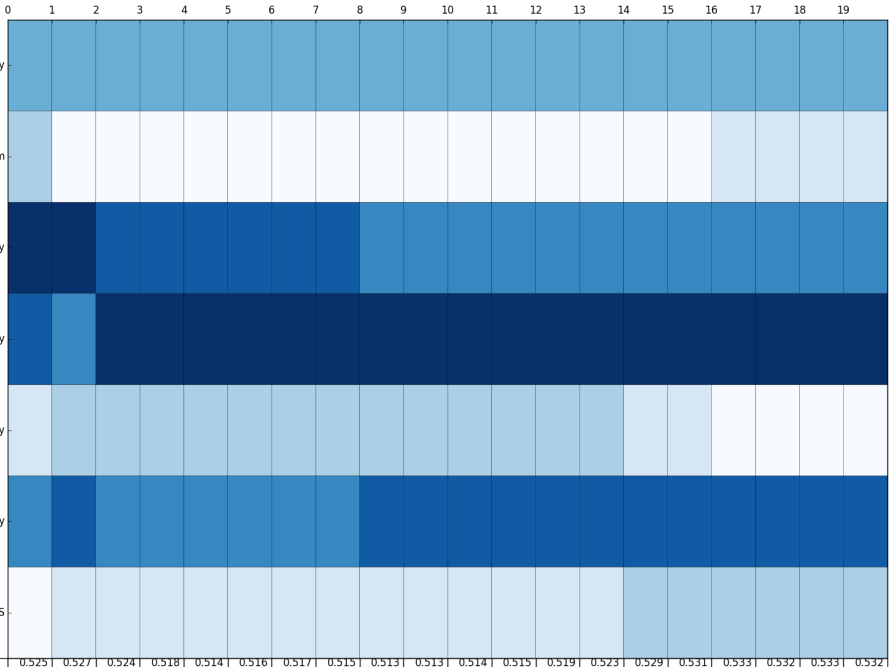
Soft Hybrid - Certainty	0.129	0.113	0.112	0.133	0.159	0.175	0.185	0.194	0.209	0.22	0.227	0.236	0.233	0.237	0.245	0.246	0.246	0.249	0.25	0.25
Random	0.132	0.153	0.169	0.196	0.206	0.205	0.219	0.218	0.229	0.238	0.245	0.244	0.244	0.245	0.247	0.247	0.251	0.254	0.256	0.261
MDS - Certainty	0.102	0.153	0.146	0.147	0.169	0.182	0.194	0.209	0.22	0.23	0.233	0.234	0.238	0.244	0.247	0.251	0.251	0.254	0.253	0.254
MDS - Uncertainty	0.104	0.145	0.142	0.16	0.186	0.195	0.195	0.21	0.22	0.223	0.231	0.237	0.244	0.251	0.253	0.259	0.26	0.258	0.26	0.262
Hard Hybrid - Certainty	0.098	0.121	0.164	0.183	0.208	0.221	0.227	0.235	0.238	0.238	0.235	0.234	0.236	0.245	0.245	0.251	0.255	0.255	0.255	0.255
Hard Hybrid - Uncertainty	0.051	0.088	0.11	0.15	0.187	0.195	0.2	0.209	0.219	0.226	0.229	0.228	0.229	0.228	0.227	0.233	0.237	0.237	0.237	0.239
BCS	0.1	0.137	0.137	0.16	0.166	0.187	0.199	0.204	0.212	0.219	0.22	0.225	0.229	0.229	0.231	0.234	0.238	0.237	0.239	0.239

BFMTV_CultureEtVous_2012-02-15_065040, Head Track: 180, Speaker Track: 34, OCR Name: 4 HvH 67X67 SvS 27X27 SvH 113



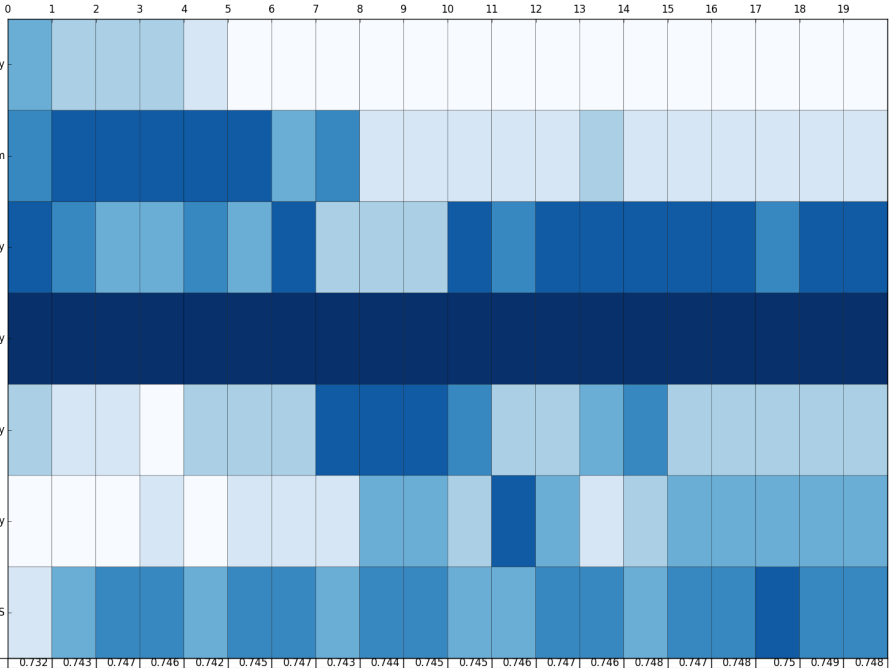
Soft Hybrid - Certainty	0.246	0.235	0.232	0.235	0.241	0.259	0.272	0.282	0.281	0.287	0.295	0.295	0.298	0.299	0.3	0.301	0.301	0.305	0.309	0.31
Random	0.235	0.235	0.239	0.229	0.242	0.242	0.258	0.265	0.272	0.282	0.286	0.291	0.296	0.299	0.298	0.303	0.301	0.303	0.305	0.305
MDS - Certainty	0.211	0.244	0.254	0.256	0.271	0.288	0.286	0.292	0.3	0.304	0.303	0.307	0.309	0.313	0.313	0.313	0.313	0.314	0.315	0.318
MDS - Uncertainty	0.208	0.248	0.248	0.262	0.259	0.266	0.269	0.276	0.279	0.284	0.292	0.294	0.296	0.298	0.3	0.301	0.303	0.306	0.307	0.308
Hard Hybrid - Certainty	0.249	0.258	0.26	0.281	0.286	0.294	0.297	0.295	0.303	0.303	0.31	0.313	0.317	0.321	0.321	0.324	0.325	0.325	0.327	0.329
Hard Hybrid - Uncertainty	0.207	0.226	0.235	0.232	0.24	0.26	0.269	0.276	0.282	0.289	0.289	0.288	0.293	0.295	0.296	0.298	0.3	0.303	0.304	0.308
BCS	0.249	0.244	0.241	0.245	0.247	0.263	0.277	0.289	0.293	0.297	0.304	0.305	0.312	0.316	0.32	0.323	0.325	0.328	0.33	0.332

LCP_CaVousRegarde_2011-12-20_204600, Head Track: 397, Speaker Track: 332, OCR Name: 30 HvH 282X282 SvS 268X268 SvH 406



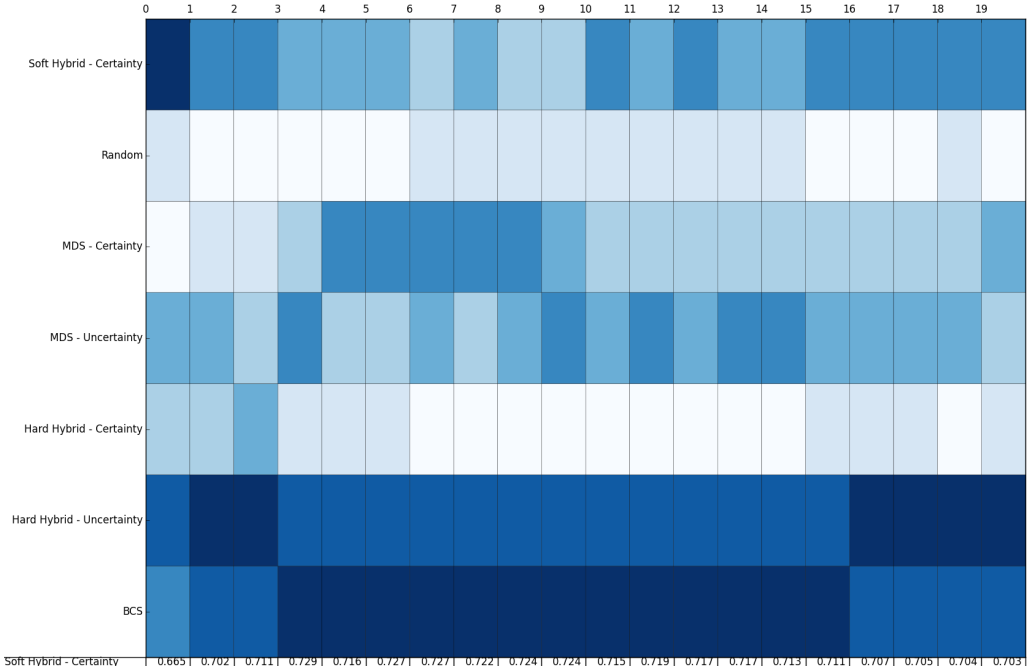
Soft Hybrid - Certainty	0.525	0.527	0.524	0.518	0.514	0.516	0.517	0.515	0.513	0.513	0.514	0.515	0.519	0.523	0.529	0.531	0.533	0.532	0.533	0.532
Random	0.526	0.568	0.566	0.565	0.565	0.556	0.555	0.547	0.55	0.55	0.549	0.549	0.546	0.546	0.547	0.545	0.542	0.54	0.54	0.54
MDS - Certainty	0.44	0.464	0.486	0.493	0.493	0.498	0.489	0.492	0.499	0.501	0.504	0.504	0.503	0.507	0.505	0.505	0.505	0.506	0.505	0.505
MDS - Uncertainty	0.483	0.489	0.465	0.467	0.475	0.479	0.482	0.479	0.475	0.481	0.487	0.489	0.494	0.492	0.491	0.492	0.491	0.493	0.493	0.492
Hard Hybrid - Certainty	0.528	0.53	0.53	0.533	0.527	0.529	0.528	0.525	0.52	0.521	0.528	0.535	0.533	0.539	0.539	0.54	0.542	0.545	0.544	0.542
Hard Hybrid - Uncertainty	0.497	0.484	0.514	0.51	0.502	0.503	0.495	0.495	0.495	0.497	0.494	0.498	0.497	0.498	0.496	0.497	0.501	0.501	0.502	0.5
BCS	0.568	0.563	0.548	0.554	0.548	0.534	0.54	0.547	0.549	0.547	0.544	0.544	0.546	0.54	0.537	0.535	0.537	0.536	0.537	0.535

LCP_CaVousRegarde_2012-01-19_204700, Head Track: 371, Speaker Track: 290, OCR Name: 22 HvH 244X244 SvS 229X229 SvH 384



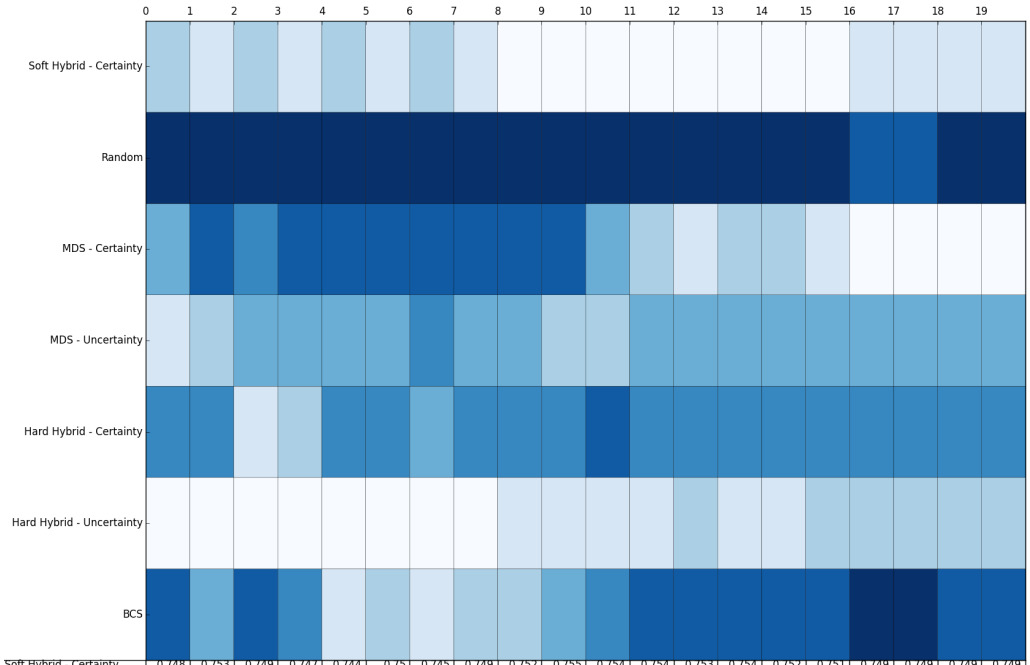
Soft Hybrid - Certainty	0.732	0.743	0.747	0.746	0.742	0.745	0.747	0.743	0.744	0.745	0.745	0.746	0.747	0.746	0.748	0.747	0.748	0.75	0.749	0.748
Random	0.709	0.704	0.702	0.714	0.723	0.726	0.737	0.735	0.74	0.743	0.744	0.743	0.746	0.744	0.744	0.745	0.746	0.747	0.746	0.745
MDS - Certainty	0.694	0.726	0.733	0.739	0.732	0.735	0.731	0.739	0.739	0.739	0.736	0.74	0.738	0.737	0.739	0.737	0.739	0.741	0.739	0.737
MDS - Uncertainty	0.651	0.686	0.687	0.681	0.689	0.699	0.699	0.704	0.71	0.714	0.714	0.715	0.719	0.717	0.717	0.719	0.717	0.718	0.716	0.718
Hard Hybrid - Certainty	0.743	0.746	0.748	0.75	0.74	0.741	0.742	0.734	0.733	0.736	0.738	0.741	0.742	0.741	0.742	0.743	0.744	0.743	0.744	0.744
Hard Hybrid - Uncertainty	0.78	0.753	0.75	0.75	0.743	0.745	0.746	0.74	0.737	0.738	0.742	0.738	0.742	0.744	0.743	0.742	0.743	0.742	0.742	0.744
BCS	0.751	0.731	0.726	0.734	0.736	0.735	0.736	0.736	0.736	0.737	0.74	0.741	0.74	0.74	0.743	0.741	0.743	0.741	0.742	0.742

LCP_CaVousRegarde_2012-01-25_204700, Head Track: 428, Speaker Track: 352, OCR Name: 25 HvH 281X281 SvS 282X282 SvH 429



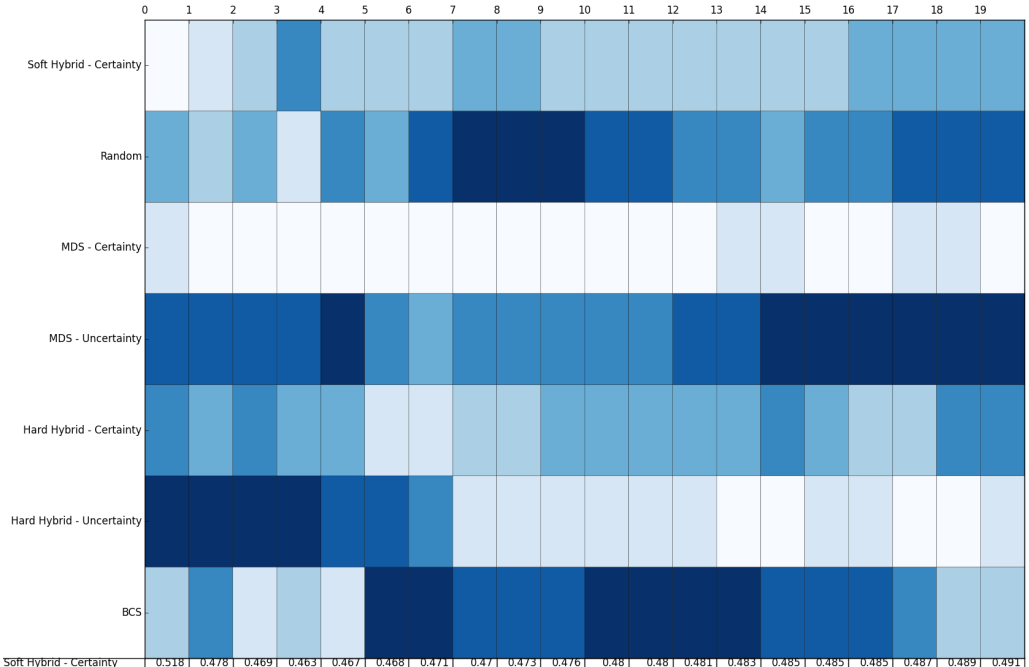
Soft Hybrid - Certainty	0.665	0.702	0.711	0.729	0.716	0.727	0.727	0.722	0.724	0.724	0.715	0.719	0.717	0.717	0.713	0.711	0.707	0.705	0.704	0.703
Random	0.746	0.761	0.76	0.755	0.757	0.746	0.739	0.743	0.735	0.736	0.733	0.731	0.731	0.732	0.733	0.733	0.73	0.727	0.724	0.726
MDS - Certainty	0.756	0.746	0.736	0.73	0.712	0.72	0.718	0.715	0.72	0.722	0.723	0.725	0.722	0.72	0.719	0.719	0.719	0.717	0.718	0.717
MDS - Uncertainty	0.722	0.723	0.729	0.727	0.725	0.732	0.724	0.725	0.722	0.718	0.718	0.718	0.717	0.713	0.71	0.713	0.716	0.714	0.716	0.718
Hard Hybrid - Certainty	0.744	0.729	0.728	0.738	0.742	0.741	0.741	0.744	0.736	0.738	0.737	0.734	0.733	0.732	0.734	0.728	0.726	0.724	0.724	0.722
Hard Hybrid - Uncertainty	0.682	0.686	0.698	0.706	0.699	0.702	0.702	0.694	0.697	0.696	0.695	0.697	0.69	0.689	0.689	0.685	0.681	0.678	0.677	0.676
BCS	0.698	0.701	0.706	0.701	0.69	0.686	0.682	0.687	0.693	0.695	0.695	0.687	0.685	0.684	0.684	0.684	0.683	0.682	0.684	0.681

LCP_EntreLesLignes_2011-12-16_192700, Head Track: 560, Speaker Track: 253, OCR Name: 15 HvH 345X345 SvS 203X203 SvH 493



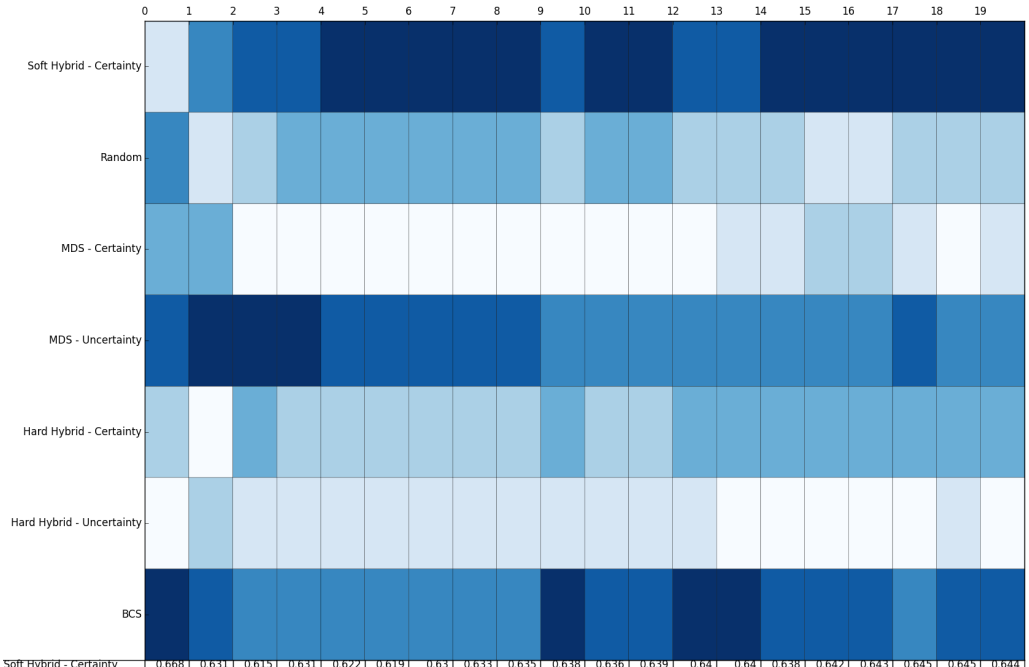
Soft Hybrid - Certainty	0.748	0.753	0.749	0.747	0.744	0.75	0.745	0.749	0.752	0.755	0.754	0.754	0.753	0.754	0.752	0.751	0.749	0.749	0.749	0.749
Random	0.708	0.717	0.711	0.719	0.729	0.73	0.731	0.731	0.733	0.733	0.733	0.733	0.735	0.738	0.739	0.738	0.739	0.739	0.738	0.739
MDS - Certainty	0.74	0.738	0.737	0.737	0.732	0.736	0.735	0.739	0.735	0.737	0.741	0.744	0.744	0.743	0.745	0.748	0.75	0.75	0.75	0.751
MDS - Uncertainty	0.75	0.742	0.741	0.745	0.743	0.742	0.738	0.74	0.741	0.743	0.743	0.741	0.742	0.742	0.742	0.741	0.741	0.741	0.743	0.743
Hard Hybrid - Certainty	0.716	0.738	0.75	0.746	0.74	0.738	0.739	0.739	0.741	0.741	0.739	0.738	0.74	0.741	0.74	0.739	0.74	0.739	0.74	0.741
Hard Hybrid - Uncertainty	0.791	0.785	0.77	0.765	0.758	0.761	0.76	0.755	0.751	0.749	0.747	0.746	0.743	0.745	0.746	0.746	0.746	0.746	0.747	0.746
BCS	0.71	0.741	0.735	0.741	0.747	0.747	0.748	0.745	0.742	0.742	0.74	0.738	0.74	0.739	0.74	0.739	0.739	0.739	0.739	0.741

LCP_EntreLesLignes_2012-01-27_192700, Head Track: 599, Speaker Track: 267, OCR Name: 14 HvH 329X329 SvS 210X210 SvH 460



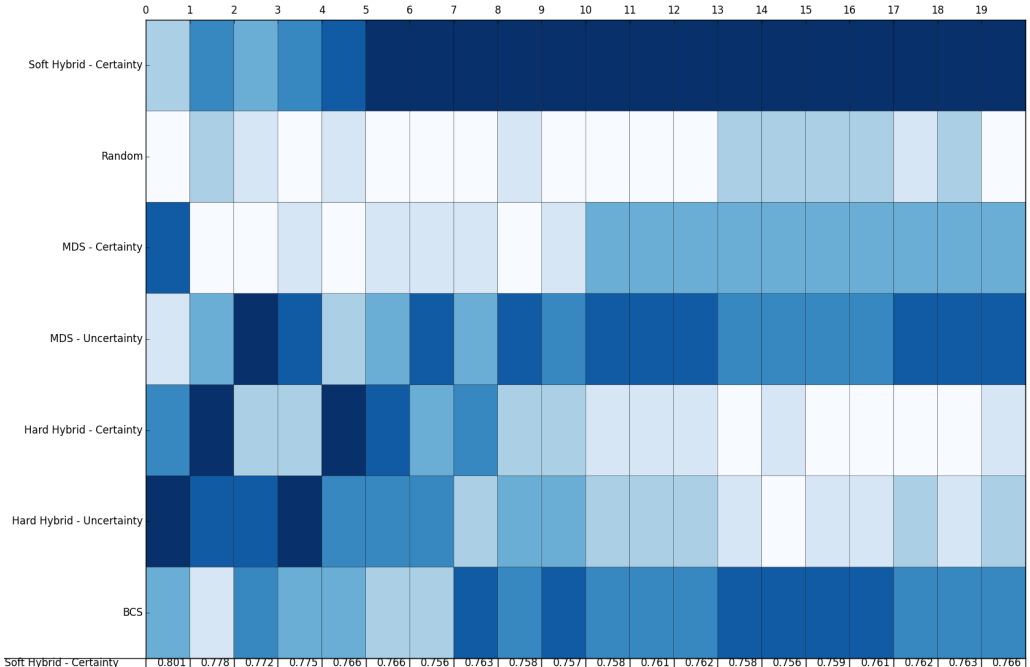
Soft Hybrid - Certainty	0.518	0.478	0.469	0.463	0.467	0.468	0.471	0.47	0.473	0.476	0.48	0.48	0.481	0.483	0.485	0.485	0.485	0.487	0.489	0.491
Random	0.466	0.474	0.468	0.471	0.463	0.462	0.459	0.456	0.457	0.461	0.469	0.471	0.475	0.476	0.48	0.482	0.481	0.484	0.488	0.489
MDS - Certainty	0.496	0.507	0.479	0.484	0.479	0.488	0.489	0.489	0.492	0.498	0.495	0.495	0.496	0.495	0.493	0.495	0.496	0.496	0.497	0.502
MDS - Uncertainty	0.455	0.461	0.452	0.452	0.454	0.46	0.466	0.465	0.468	0.468	0.47	0.472	0.472	0.473	0.474	0.477	0.478	0.478	0.479	0.482
Hard Hybrid - Certainty	0.456	0.465	0.464	0.464	0.463	0.474	0.473	0.471	0.475	0.471	0.473	0.473	0.479	0.48	0.479	0.485	0.486	0.489	0.489	0.491
Hard Hybrid - Uncertainty	0.427	0.454	0.443	0.451	0.455	0.459	0.466	0.472	0.477	0.48	0.484	0.49	0.492	0.496	0.494	0.494	0.495	0.497	0.499	0.5
BCS	0.467	0.465	0.475	0.47	0.469	0.459	0.458	0.459	0.467	0.467	0.469	0.468	0.469	0.472	0.476	0.479	0.48	0.487	0.49	0.492

LCP_EntreLesLignes_2012-05-11_192200, Head Track: 690, Speaker Track: 305, OCR Name: 13 HvH 310X310 SvS 240X240 SvH 596



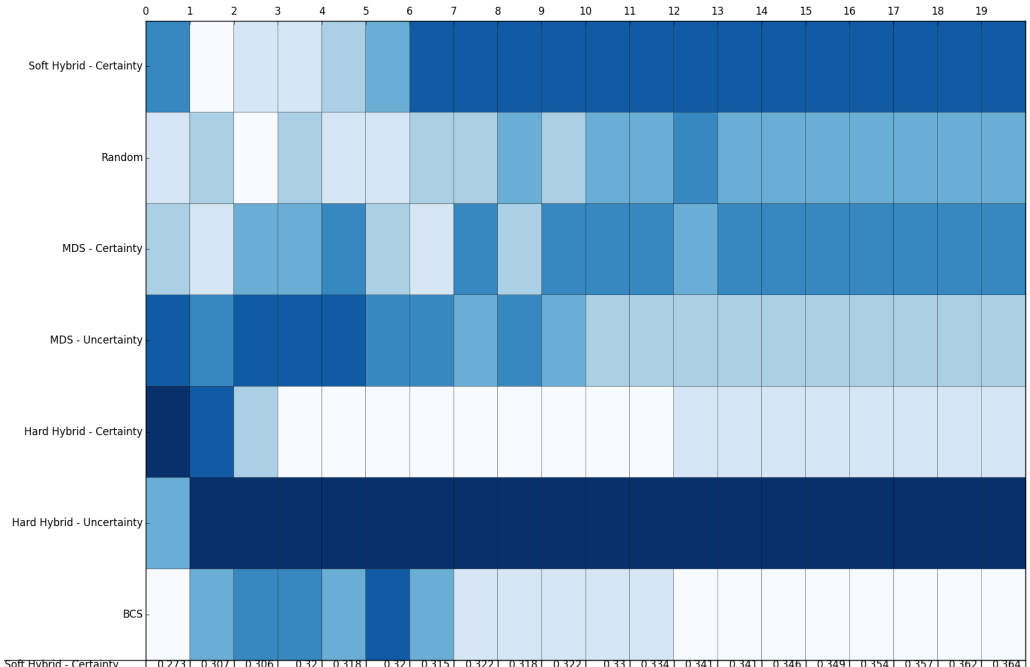
Soft Hybrid - Certainty	0.668	0.631	0.615	0.631	0.622	0.619	0.63	0.633	0.635	0.638	0.636	0.639	0.64	0.64	0.638	0.642	0.643	0.645	0.645	0.644
Random	0.644	0.666	0.656	0.648	0.65	0.65	0.649	0.651	0.655	0.655	0.655	0.657	0.655	0.656	0.658	0.66	0.662	0.655	0.656	0.657
MDS - Certainty	0.649	0.658	0.674	0.676	0.672	0.672	0.674	0.671	0.669	0.671	0.669	0.669	0.667	0.663	0.661	0.66	0.659	0.658	0.66	0.659
MDS - Uncertainty	0.641	0.615	0.608	0.623	0.622	0.627	0.63	0.635	0.636	0.643	0.646	0.643	0.646	0.648	0.648	0.65	0.647	0.645	0.648	0.648
Hard Hybrid - Certainty	0.666	0.668	0.655	0.654	0.651	0.65	0.649	0.651	0.656	0.654	0.656	0.659	0.654	0.653	0.653	0.654	0.652	0.65	0.651	0.649
Hard Hybrid - Uncertainty	0.696	0.659	0.667	0.667	0.662	0.656	0.659	0.658	0.666	0.669	0.666	0.661	0.663	0.664	0.664	0.661	0.663	0.662	0.659	0.659
BCS	0.61	0.629	0.637	0.64	0.638	0.632	0.636	0.636	0.639	0.636	0.639	0.641	0.639	0.639	0.644	0.645	0.644	0.648	0.648	0.647

LCP_LCPInfo13h30_2012-01-24_132700, Head Track: 212, Speaker Track: 163, OCR Name: 13 HvH 128X128 SvS 130X130 SvH 179



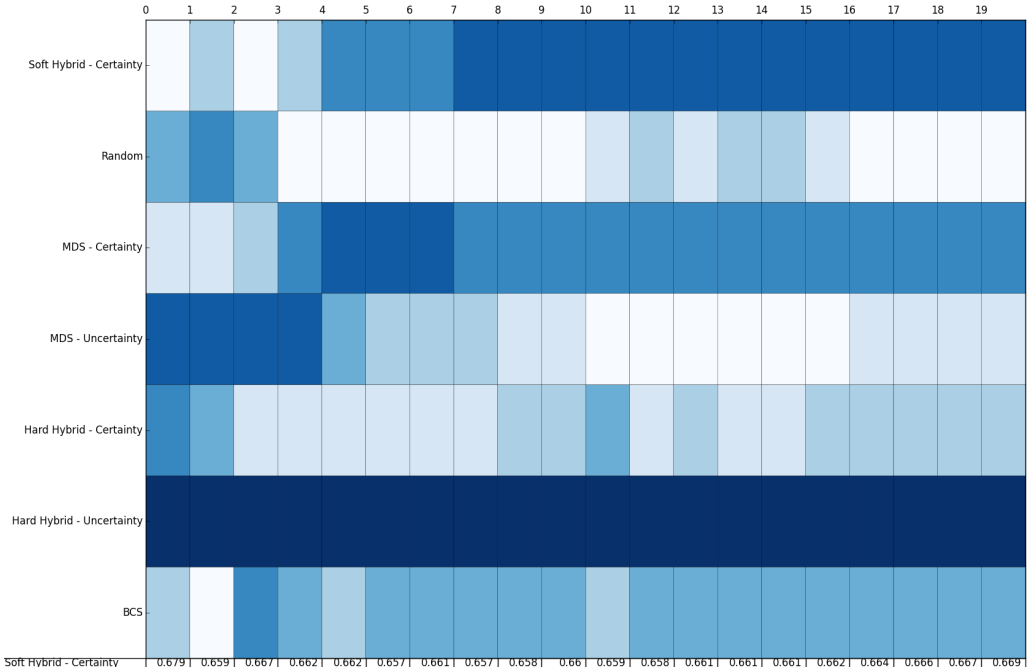
Soft Hybrid - Certainty	0.801	0.778	0.772	0.775	0.766	0.766	0.756	0.763	0.758	0.757	0.758	0.761	0.762	0.758	0.756	0.759	0.761	0.762	0.763	0.766
Random	0.807	0.785	0.783	0.799	0.794	0.801	0.801	0.802	0.795	0.797	0.799	0.802	0.798	0.792	0.792	0.794	0.795	0.795	0.794	0.797
MDS - Certainty	0.771	0.79	0.792	0.787	0.796	0.797	0.794	0.795	0.796	0.794	0.785	0.786	0.787	0.785	0.788	0.787	0.787	0.783	0.784	0.78
MDS - Uncertainty	0.803	0.781	0.764	0.763	0.779	0.774	0.765	0.775	0.769	0.775	0.775	0.772	0.776	0.777	0.776	0.776	0.775	0.774	0.773	0.772
Hard Hybrid - Certainty	0.773	0.768	0.777	0.778	0.762	0.768	0.772	0.775	0.782	0.788	0.789	0.793	0.793	0.792	0.794	0.796	0.796	0.799	0.8	0.796
Hard Hybrid - Uncertainty	0.747	0.774	0.766	0.755	0.772	0.77	0.767	0.778	0.781	0.786	0.787	0.789	0.793	0.792	0.796	0.794	0.796	0.791	0.795	0.795
BCS	0.782	0.788	0.769	0.776	0.779	0.779	0.778	0.764	0.771	0.774	0.779	0.781	0.779	0.773	0.771	0.771	0.775	0.777	0.777	0.779

LCP_LCPInfo13h30_2012-01-25_132700, Head Track: 331, Speaker Track: 172, OCR Name: 10 HvH 208X208 SvS 145X145 SvH 285



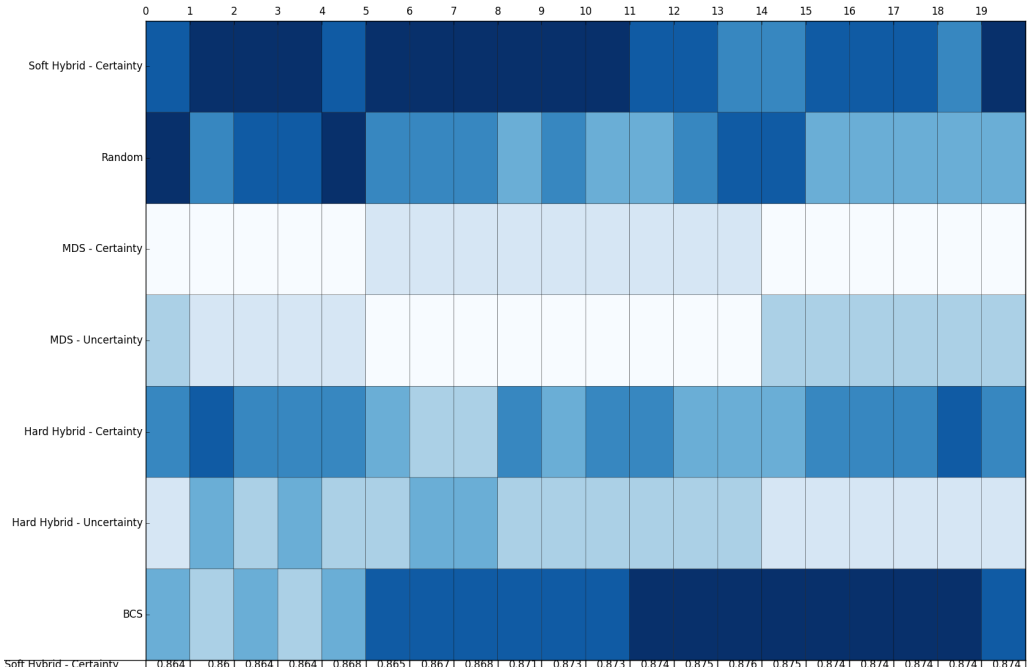
Soft Hybrid - Certainty	0.273	0.307	0.306	0.32	0.318	0.32	0.315	0.322	0.318	0.322	0.33	0.334	0.341	0.341	0.346	0.349	0.354	0.357	0.362	0.364
Random	0.284	0.289	0.318	0.32	0.32	0.325	0.332	0.343	0.353	0.36	0.362	0.36	0.359	0.365	0.377	0.381	0.382	0.385	0.389	0.395
MDS - Certainty	0.282	0.295	0.303	0.314	0.31	0.322	0.333	0.34	0.354	0.354	0.358	0.358	0.362	0.362	0.367	0.373	0.379	0.383	0.387	0.391
MDS - Uncertainty	0.269	0.285	0.282	0.287	0.307	0.316	0.324	0.341	0.351	0.357	0.363	0.371	0.376	0.38	0.383	0.388	0.389	0.396	0.401	0.407
Hard Hybrid - Certainty	0.265	0.276	0.305	0.331	0.34	0.346	0.355	0.363	0.371	0.382	0.395	0.396	0.399	0.405	0.413	0.419	0.425	0.429	0.431	0.431
Hard Hybrid - Uncertainty	0.274	0.276	0.282	0.285	0.288	0.302	0.298	0.301	0.305	0.312	0.313	0.32	0.324	0.325	0.332	0.343	0.341	0.344	0.349	0.35
BCS	0.295	0.286	0.289	0.3	0.312	0.314	0.329	0.349	0.356	0.365	0.38	0.392	0.4	0.408	0.419	0.428	0.437	0.444	0.451	0.455

LCP_LCPInfo13h30_2012-01-27_132700, Head Track: 306, Speaker Track: 177, OCR Name: 11 HvH 182X182 SvS 135X135 SvH 270



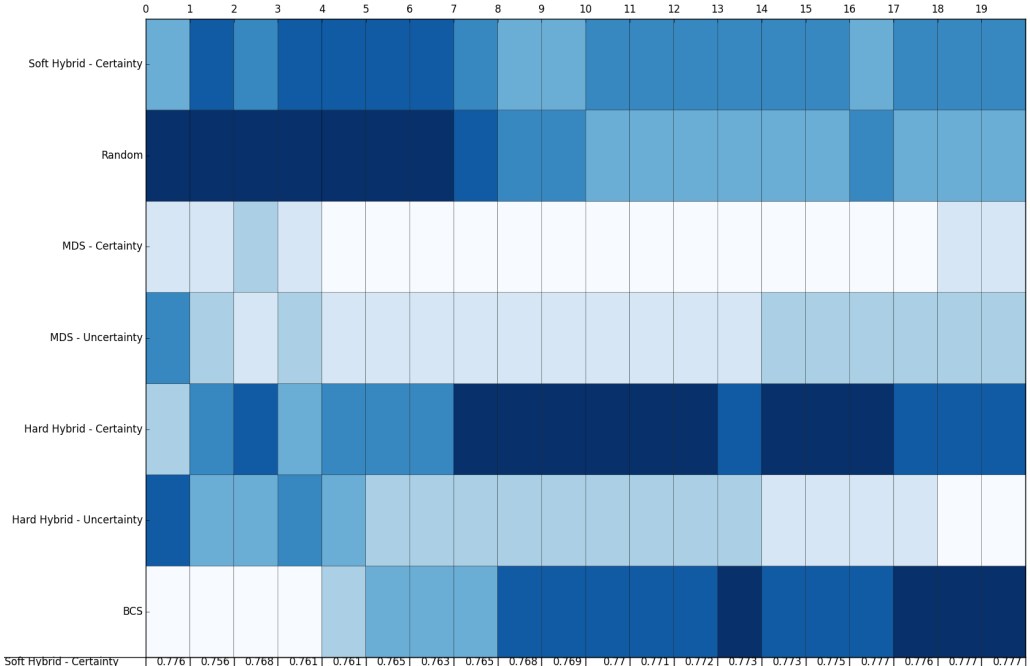
Soft Hybrid - Certainty	0.679	0.659	0.667	0.662	0.662	0.657	0.661	0.657	0.658	0.66	0.659	0.658	0.661	0.661	0.661	0.662	0.664	0.666	0.667	0.669
Random	0.655	0.656	0.664	0.676	0.67	0.678	0.681	0.679	0.675	0.674	0.673	0.671	0.674	0.672	0.675	0.677	0.678	0.679	0.682	0.684
MDS - Certainty	0.679	0.66	0.667	0.656	0.653	0.657	0.661	0.659	0.659	0.666	0.666	0.666	0.666	0.667	0.666	0.666	0.667	0.669	0.669	0.671
MDS - Uncertainty	0.645	0.63	0.649	0.653	0.663	0.666	0.672	0.673	0.672	0.672	0.676	0.677	0.676	0.679	0.678	0.677	0.676	0.677	0.68	0.68
Hard Hybrid - Certainty	0.649	0.657	0.667	0.669	0.669	0.674	0.675	0.674	0.672	0.672	0.669	0.672	0.671	0.676	0.677	0.676	0.676	0.676	0.677	0.677
Hard Hybrid - Uncertainty	0.617	0.622	0.622	0.623	0.635	0.639	0.643	0.645	0.649	0.651	0.649	0.649	0.651	0.652	0.655	0.658	0.66	0.663	0.664	0.667
BCS	0.672	0.671	0.661	0.662	0.667	0.662	0.671	0.668	0.667	0.67	0.672	0.671	0.67	0.671	0.673	0.672	0.673	0.67	0.672	0.672

LCP_PileEtFace_2011-11-19_085700, Head Track: 243, Speaker Track: 211, OCR Name: 40 HvH 173X173 SvS 162X162 SvH 289



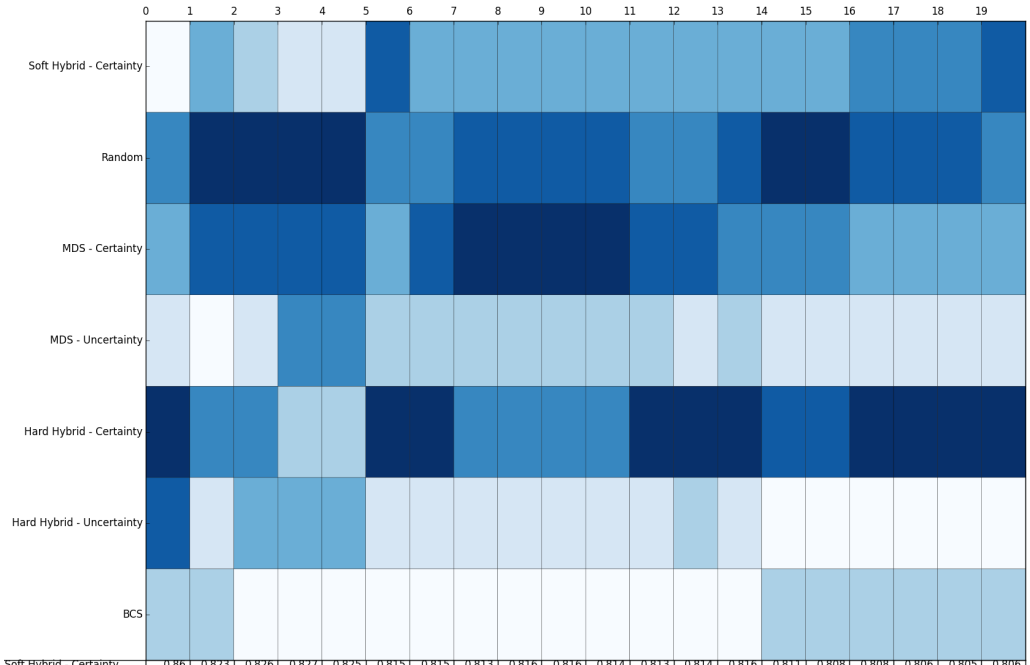
Soft Hybrid - Certainty	0.864	0.86	0.864	0.864	0.868	0.865	0.867	0.868	0.871	0.873	0.873	0.874	0.875	0.876	0.875	0.874	0.874	0.874	0.874	0.874
Random	0.863	0.87	0.868	0.868	0.868	0.873	0.874	0.876	0.877	0.877	0.879	0.877	0.877	0.875	0.875	0.876	0.876	0.876	0.877	0.877
MDS - Certainty	0.894	0.896	0.896	0.891	0.89	0.891	0.89	0.882	0.884	0.884	0.885	0.885	0.885	0.885	0.887	0.887	0.887	0.887	0.886	0.887
MDS - Uncertainty	0.876	0.882	0.886	0.888	0.888	0.892	0.892	0.89	0.885	0.885	0.886	0.885	0.885	0.886	0.88	0.879	0.877	0.878	0.879	0.879
Hard Hybrid - Certainty	0.866	0.866	0.871	0.871	0.875	0.876	0.878	0.876	0.875	0.877	0.876	0.876	0.878	0.877	0.876	0.875	0.875	0.875	0.874	0.876
Hard Hybrid - Uncertainty	0.878	0.874	0.876	0.875	0.877	0.88	0.877	0.876	0.88	0.88	0.882	0.878	0.879	0.881	0.881	0.882	0.881	0.881	0.882	0.883
BCS	0.875	0.876	0.876	0.877	0.876	0.872	0.873	0.872	0.873	0.874	0.875	0.874	0.873	0.872	0.873	0.874	0.873	0.873	0.874	0.875

LCP_PileEtFace_2011-12-01_192700, Head Track: 389, Speaker Track: 255, OCR Name: 51 HvH 250X250 SvS 197X197 SvH 349



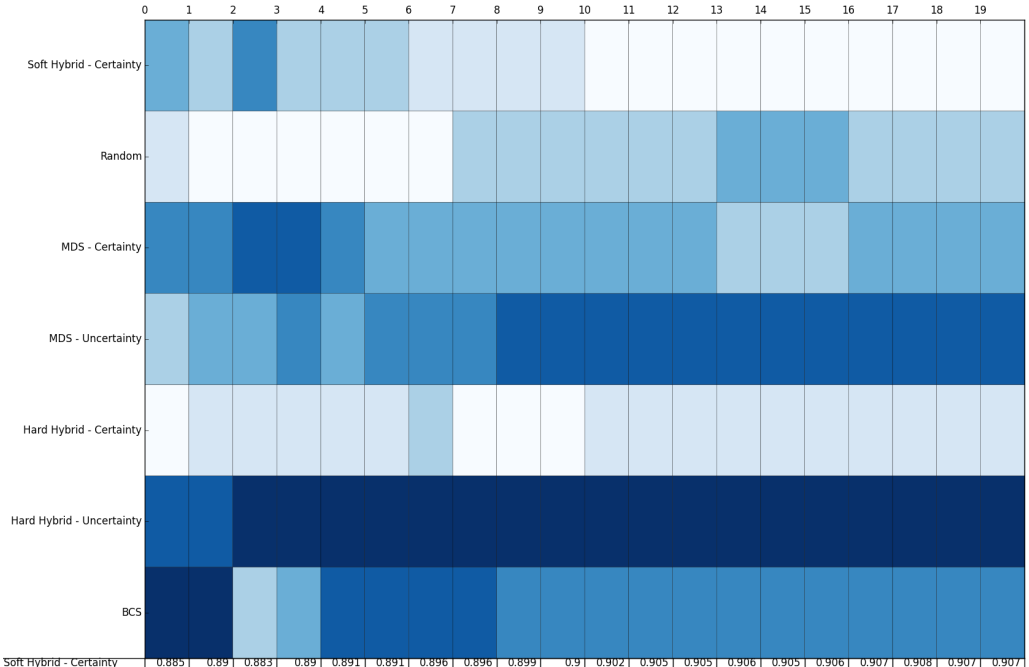
Soft Hybrid - Certainty	0.776	0.756	0.768	0.761	0.761	0.765	0.763	0.765	0.768	0.769	0.77	0.771	0.772	0.773	0.773	0.775	0.777	0.776	0.777	0.777
Random	0.762	0.752	0.761	0.756	0.756	0.758	0.763	0.765	0.768	0.769	0.772	0.773	0.774	0.775	0.776	0.776	0.776	0.777	0.779	0.778
MDS - Certainty	0.78	0.783	0.783	0.78	0.779	0.785	0.786	0.784	0.787	0.79	0.791	0.786	0.786	0.787	0.789	0.789	0.789	0.789	0.789	0.789
MDS - Uncertainty	0.772	0.778	0.785	0.779	0.777	0.783	0.781	0.778	0.78	0.781	0.783	0.784	0.784	0.783	0.783	0.783	0.783	0.782	0.782	0.782
Hard Hybrid - Certainty	0.777	0.764	0.766	0.766	0.768	0.766	0.764	0.763	0.765	0.764	0.766	0.766	0.767	0.768	0.767	0.768	0.769	0.77	0.772	0.772
Hard Hybrid - Uncertainty	0.765	0.77	0.776	0.766	0.773	0.776	0.774	0.776	0.777	0.777	0.776	0.774	0.779	0.782	0.785	0.785	0.786	0.788	0.789	0.789
BCS	0.798	0.784	0.789	0.783	0.775	0.775	0.771	0.772	0.767	0.768	0.77	0.769	0.769	0.767	0.768	0.769	0.77	0.77	0.771	0.772

LCP_PileEtFace_2012-01-12_192700, Head Track: 330, Speaker Track: 269, OCR Name: 52 HvH 228X228 SvS 218X218 SvH 373



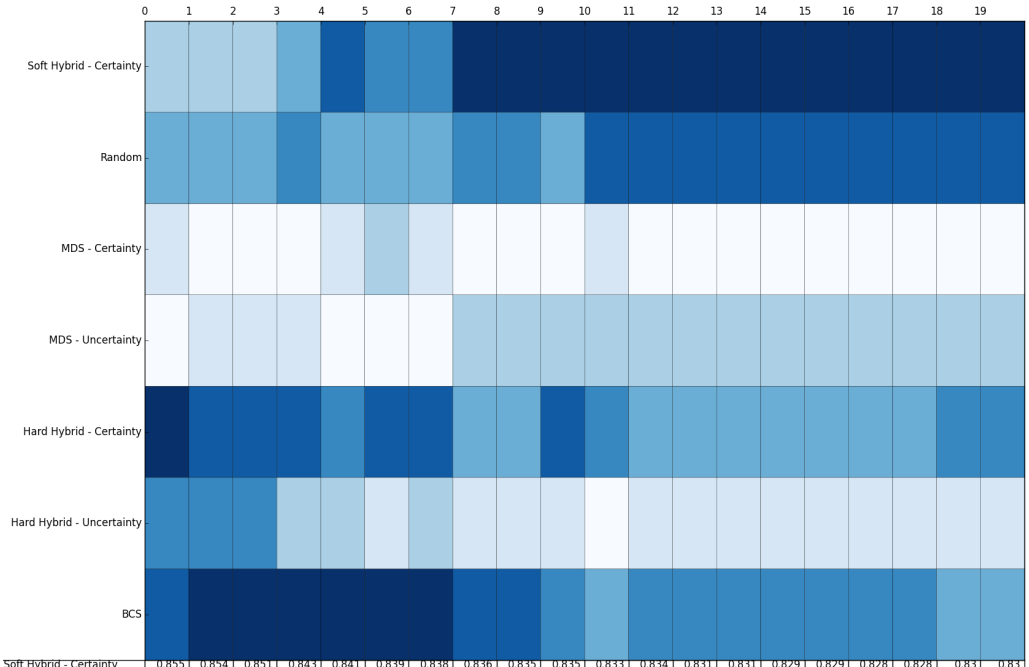
Soft Hybrid - Certainty	0.86	0.823	0.826	0.827	0.825	0.815	0.815	0.813	0.816	0.816	0.814	0.813	0.814	0.816	0.811	0.808	0.808	0.806	0.805	0.806
Random	0.827	0.794	0.798	0.807	0.811	0.817	0.812	0.809	0.806	0.808	0.806	0.807	0.806	0.803	0.802	0.8	0.802	0.805	0.804	0.808
MDS - Certainty	0.834	0.811	0.813	0.818	0.813	0.818	0.808	0.809	0.8	0.796	0.8	0.803	0.804	0.806	0.808	0.808	0.809	0.811	0.811	0.811
MDS - Uncertainty	0.854	0.837	0.834	0.822	0.819	0.82	0.818	0.818	0.822	0.821	0.824	0.825	0.827	0.826	0.825	0.824	0.823	0.825	0.822	0.823
Hard Hybrid - Certainty	0.802	0.813	0.816	0.823	0.822	0.814	0.807	0.809	0.811	0.81	0.807	0.795	0.799	0.801	0.804	0.803	0.8	0.803	0.802	0.798
Hard Hybrid - Uncertainty	0.807	0.831	0.82	0.823	0.821	0.823	0.83	0.829	0.829	0.831	0.828	0.828	0.827	0.827	0.829	0.829	0.826	0.828	0.827	0.827
BCS	0.842	0.826	0.836	0.836	0.836	0.826	0.833	0.835	0.837	0.838	0.835	0.833	0.834	0.829	0.822	0.822	0.82	0.816	0.815	0.814

LCP_PileEtFace_2012-01-19_192700, Head Track: 302, Speaker Track: 246, OCR Name: 43 HvH 222X222 SvS 191X191 SvH 290



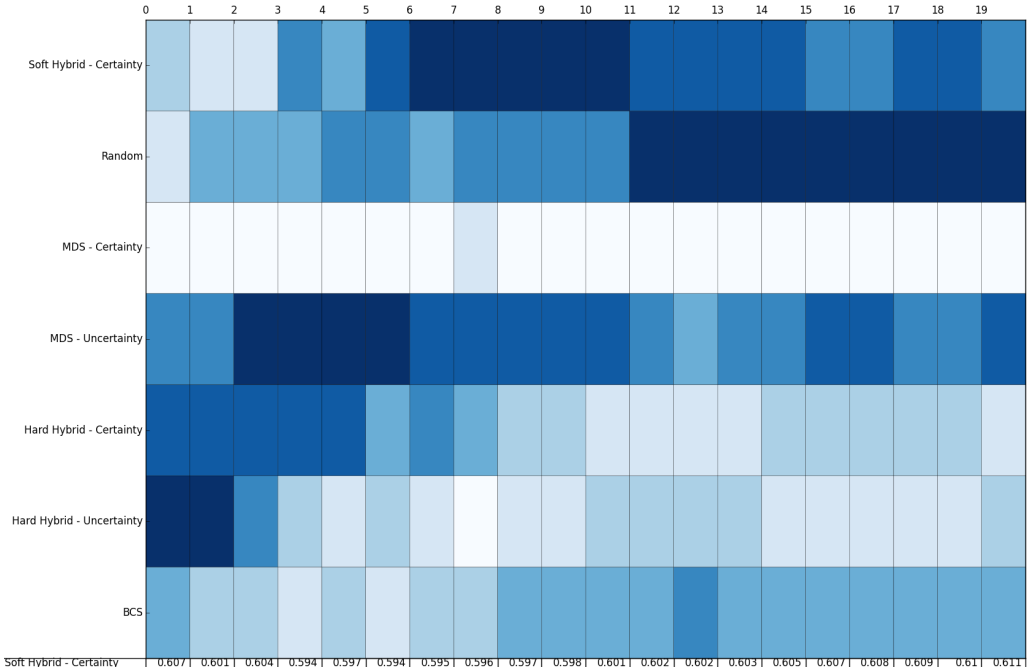
Soft Hybrid - Certainty	0.885	0.89	0.883	0.89	0.891	0.891	0.896	0.896	0.899	0.9	0.902	0.905	0.905	0.906	0.905	0.906	0.907	0.908	0.907	0.907
Random	0.897	0.902	0.902	0.901	0.899	0.899	0.898	0.896	0.898	0.9	0.899	0.9	0.898	0.896	0.897	0.897	0.899	0.9	0.901	0.9
MDS - Certainty	0.875	0.879	0.881	0.882	0.884	0.889	0.888	0.892	0.895	0.895	0.895	0.897	0.897	0.898	0.898	0.898	0.897	0.898	0.898	0.899
MDS - Uncertainty	0.886	0.882	0.884	0.883	0.886	0.888	0.887	0.888	0.887	0.888	0.888	0.887	0.888	0.888	0.888	0.888	0.888	0.89	0.891	0.891
Hard Hybrid - Certainty	0.897	0.895	0.896	0.897	0.894	0.897	0.895	0.897	0.9	0.901	0.902	0.903	0.901	0.901	0.902	0.902	0.903	0.904	0.905	0.906
Hard Hybrid - Uncertainty	0.869	0.878	0.873	0.87	0.874	0.875	0.873	0.873	0.87	0.869	0.866	0.867	0.867	0.867	0.868	0.871	0.872	0.873	0.874	0.876
BCS	0.863	0.878	0.885	0.886	0.883	0.882	0.885	0.887	0.889	0.89	0.89	0.891	0.892	0.892	0.892	0.893	0.893	0.892	0.893	0.893

LCP_PileEtFace_2012-01-26_192700, Head Track: 380, Speaker Track: 273, OCR Name: 50 HvH 223X223 SvS 207X207 SvH 349



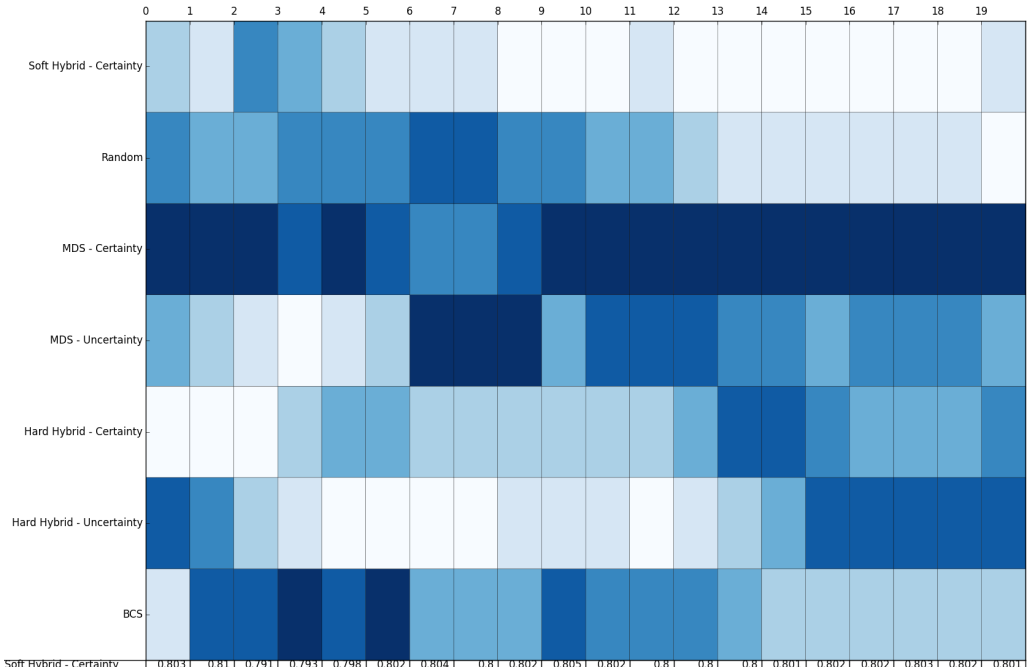
Soft Hybrid - Certainty	0.855	0.854	0.851	0.843	0.841	0.839	0.838	0.836	0.835	0.835	0.833	0.834	0.831	0.831	0.829	0.829	0.828	0.828	0.83	0.83
Random	0.853	0.851	0.845	0.843	0.846	0.844	0.843	0.838	0.84	0.841	0.836	0.836	0.835	0.833	0.833	0.834	0.832	0.832	0.832	0.832
MDS - Certainty	0.859	0.864	0.857	0.855	0.855	0.851	0.852	0.852	0.85	0.85	0.849	0.849	0.849	0.848	0.848	0.848	0.847	0.848	0.846	0.845
MDS - Uncertainty	0.863	0.854	0.852	0.853	0.855	0.855	0.853	0.847	0.846	0.841	0.842	0.842	0.84	0.838	0.839	0.84	0.841	0.84	0.841	0.84
Hard Hybrid - Certainty	0.834	0.844	0.84	0.841	0.843	0.835	0.837	0.838	0.841	0.838	0.84	0.84	0.84	0.837	0.839	0.84	0.84	0.837	0.837	0.836
Hard Hybrid - Uncertainty	0.851	0.849	0.842	0.844	0.85	0.852	0.848	0.849	0.849	0.849	0.849	0.846	0.841	0.839	0.84	0.842	0.843	0.844	0.843	0.842
BCS	0.839	0.839	0.831	0.829	0.831	0.835	0.835	0.837	0.84	0.841	0.841	0.84	0.839	0.836	0.838	0.837	0.837	0.837	0.838	0.838

LCP_TopQuestions_2012-01-25_232700, Head Track: 364, Speaker Track: 134, OCR Name: 11 HvH 204X204 SvS 102X102 SvH 310

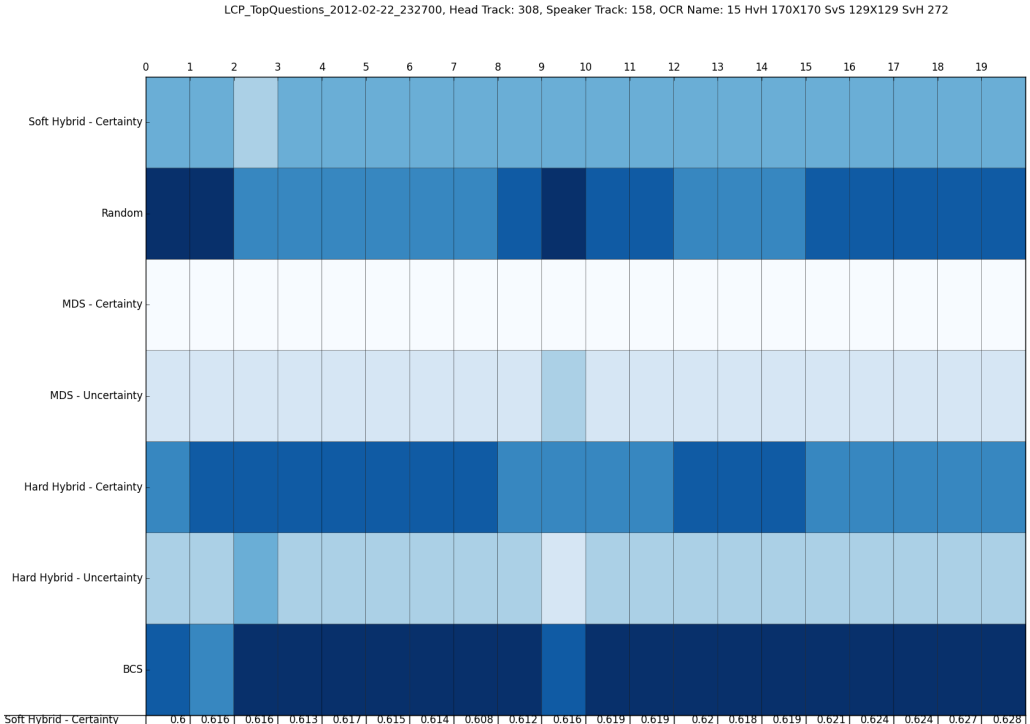


Soft Hybrid - Certainty	0.607	0.601	0.604	0.594	0.597	0.594	0.595	0.596	0.597	0.598	0.601	0.602	0.602	0.603	0.605	0.607	0.608	0.609	0.61	0.611
Random	0.608	0.589	0.598	0.596	0.596	0.597	0.603	0.599	0.603	0.605	0.602	0.598	0.595	0.596	0.594	0.595	0.598	0.599	0.598	0.599
MDS - Certainty	0.615	0.617	0.619	0.612	0.612	0.612	0.611	0.61	0.612	0.613	0.614	0.617	0.62	0.621	0.62	0.621	0.619	0.621	0.62	0.621
MDS - Uncertainty	0.586	0.588	0.575	0.574	0.584	0.585	0.595	0.596	0.599	0.598	0.601	0.603	0.605	0.604	0.607	0.606	0.607	0.609	0.61	0.609
Hard Hybrid - Certainty	0.581	0.587	0.58	0.588	0.595	0.6	0.602	0.602	0.605	0.607	0.61	0.612	0.612	0.614	0.614	0.615	0.616	0.616	0.616	0.618
Hard Hybrid - Uncertainty	0.572	0.585	0.586	0.599	0.604	0.606	0.608	0.61	0.609	0.608	0.608	0.61	0.611	0.613	0.615	0.616	0.618	0.618	0.617	0.618
BCS	0.595	0.597	0.604	0.6	0.604	0.608	0.607	0.605	0.605	0.606	0.604	0.604	0.605	0.608	0.608	0.608	0.609	0.611	0.611	0.612

LCP_TopQuestions_2012-02-14_232200, Head Track: 249, Speaker Track: 155, OCR Name: 13 HvH 133X133 SvS 127X127 SvH 272

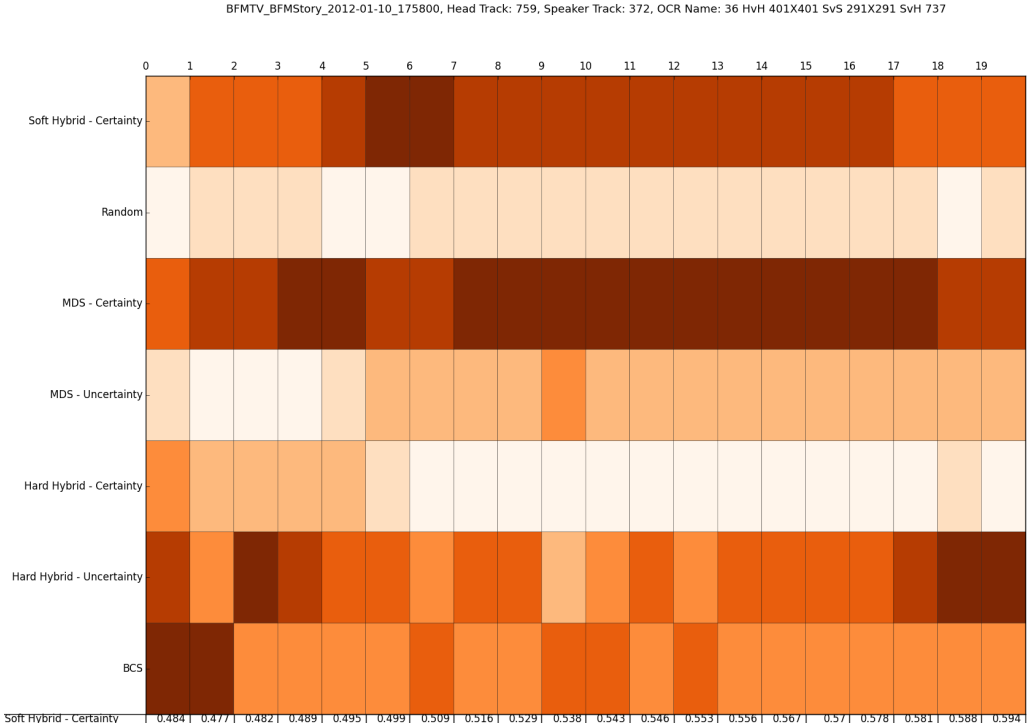


Soft Hybrid - Certainty	0.803	0.81	0.791	0.793	0.798	0.802	0.804	0.8	0.802	0.805	0.802	0.8	0.8	0.8	0.801	0.802	0.802	0.803	0.802	0.801
Random	0.795	0.794	0.792	0.789	0.794	0.793	0.791	0.792	0.793	0.795	0.796	0.799	0.798	0.799	0.8	0.8	0.801	0.802	0.801	0.802
MDS - Certainty	0.745	0.769	0.783	0.786	0.785	0.793	0.792	0.793	0.792	0.791	0.792	0.792	0.793	0.793	0.793	0.792	0.791	0.793	0.79	0.79
MDS - Uncertainty	0.798	0.796	0.802	0.805	0.803	0.798	0.79	0.789	0.791	0.795	0.793	0.792	0.794	0.796	0.795	0.795	0.794	0.794	0.795	0.795
Hard Hybrid - Certainty	0.814	0.816	0.806	0.797	0.797	0.797	0.795	0.795	0.796	0.797	0.797	0.799	0.797	0.794	0.794	0.795	0.795	0.795	0.796	0.795
Hard Hybrid - Uncertainty	0.777	0.789	0.796	0.8	0.804	0.807	0.807	0.803	0.801	0.801	0.8	0.801	0.8	0.798	0.798	0.795	0.794	0.793	0.792	0.791
BCS	0.807	0.789	0.786	0.785	0.79	0.793	0.795	0.795	0.794	0.795	0.794	0.797	0.797	0.798	0.799	0.799	0.8	0.8	0.8	0.8



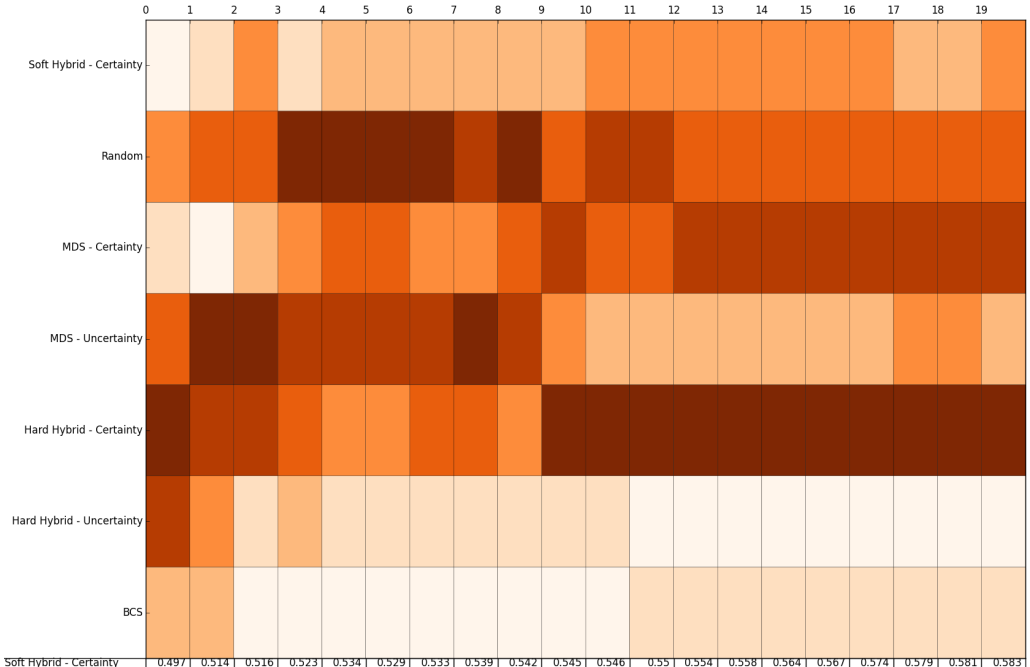
Soft Hybrid - Certainty	0.6	0.616	0.616	0.613	0.617	0.615	0.614	0.608	0.612	0.616	0.619	0.619	0.62	0.618	0.619	0.621	0.624	0.624	0.627	0.628
Random	0.541	0.556	0.585	0.592	0.596	0.594	0.594	0.594	0.599	0.594	0.597	0.602	0.606	0.608	0.611	0.611	0.613	0.616	0.617	0.619
MDS - Certainty	0.641	0.657	0.648	0.654	0.654	0.65	0.643	0.647	0.646	0.648	0.651	0.65	0.652	0.654	0.656	0.657	0.659	0.66	0.661	0.663
MDS - Uncertainty	0.625	0.62	0.632	0.63	0.628	0.634	0.632	0.63	0.63	0.63	0.637	0.637	0.638	0.64	0.642	0.643	0.646	0.647	0.647	0.649
Hard Hybrid - Certainty	0.591	0.564	0.576	0.582	0.59	0.583	0.589	0.593	0.599	0.603	0.602	0.607	0.605	0.604	0.609	0.613	0.614	0.617	0.62	0.621
Hard Hybrid - Uncertainty	0.608	0.619	0.606	0.617	0.627	0.625	0.625	0.626	0.628	0.633	0.634	0.637	0.637	0.639	0.642	0.641	0.643	0.645	0.645	0.647
BCS	0.545	0.578	0.575	0.571	0.58	0.583	0.584	0.592	0.594	0.597	0.596	0.598	0.599	0.599	0.599	0.6	0.6	0.598	0.598	0.598

Speaker score using Speaker annotation is given in orange heat maps.



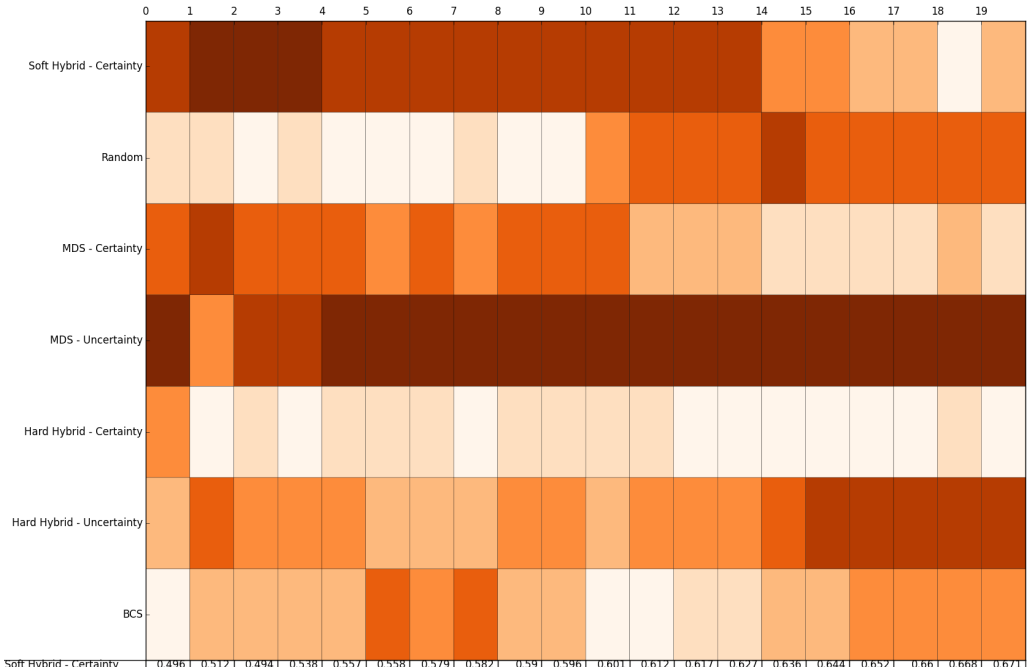
Soft Hybrid - Certainty	0.484	0.477	0.482	0.489	0.495	0.499	0.509	0.516	0.529	0.538	0.543	0.546	0.553	0.556	0.567	0.57	0.578	0.581	0.588	0.594
Random	0.544	0.535	0.548	0.553	0.568	0.561	0.557	0.555	0.571	0.568	0.577	0.582	0.583	0.593	0.593	0.597	0.603	0.609	0.613	0.614
MDS - Certainty	0.462	0.468	0.479	0.482	0.484	0.503	0.512	0.509	0.527	0.528	0.533	0.539	0.55	0.556	0.563	0.57	0.572	0.579	0.586	0.592
MDS - Uncertainty	0.533	0.554	0.55	0.556	0.547	0.54	0.541	0.55	0.56	0.561	0.566	0.575	0.578	0.586	0.591	0.596	0.602	0.606	0.608	0.611
Hard Hybrid - Certainty	0.465	0.51	0.535	0.538	0.532	0.542	0.559	0.569	0.571	0.586	0.589	0.593	0.594	0.598	0.598	0.603	0.606	0.609	0.611	0.614
Hard Hybrid - Uncertainty	0.416	0.479	0.474	0.483	0.506	0.524	0.54	0.539	0.548	0.562	0.566	0.569	0.574	0.572	0.572	0.574	0.583	0.58	0.575	0.583
BCS	0.409	0.463	0.487	0.502	0.51	0.529	0.539	0.546	0.553	0.553	0.558	0.572	0.574	0.579	0.576	0.585	0.589	0.593	0.596	0.599

BFMTV_BFMStory_2012-01-23_175800, Head Track: 782, Speaker Track: 338, OCR Name: 34 HvH 406X406 SvS 247X247 SvH 556



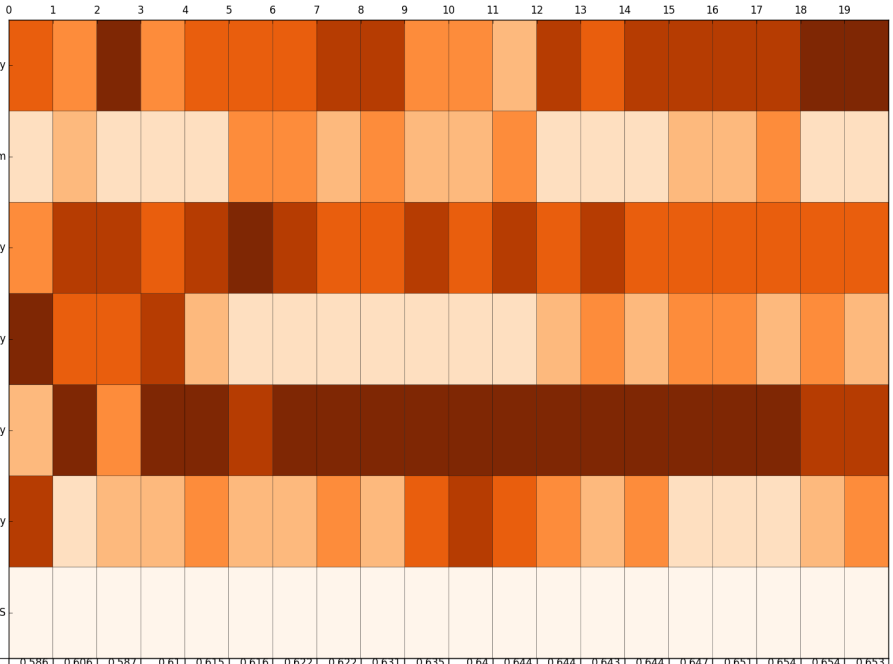
Soft Hybrid - Certainty	0.497	0.514	0.516	0.523	0.534	0.529	0.533	0.539	0.542	0.545	0.546	0.55	0.554	0.558	0.564	0.567	0.574	0.579	0.581	0.583
Random	0.47	0.495	0.503	0.49	0.501	0.513	0.517	0.525	0.528	0.537	0.54	0.543	0.547	0.551	0.554	0.559	0.564	0.565	0.571	0.574
MDS - Certainty	0.489	0.517	0.521	0.516	0.515	0.523	0.53	0.529	0.531	0.535	0.543	0.546	0.547	0.549	0.552	0.556	0.559	0.562	0.564	0.568
MDS - Uncertainty	0.464	0.453	0.491	0.51	0.513	0.522	0.519	0.523	0.53	0.541	0.548	0.55	0.555	0.56	0.567	0.572	0.574	0.576	0.579	0.583
Hard Hybrid - Certainty	0.439	0.492	0.499	0.511	0.516	0.523	0.526	0.528	0.531	0.532	0.537	0.539	0.543	0.544	0.547	0.547	0.551	0.551	0.557	0.564
Hard Hybrid - Uncertainty	0.448	0.495	0.521	0.52	0.537	0.545	0.547	0.556	0.558	0.564	0.568	0.576	0.579	0.582	0.589	0.592	0.595	0.596	0.601	0.602
BCS	0.472	0.506	0.537	0.547	0.553	0.562	0.569	0.563	0.564	0.565	0.57	0.572	0.572	0.571	0.573	0.577	0.581	0.583	0.585	0.587

BFMTV_BFMStory_2012-02-14_175800, Head Track: 753, Speaker Track: 348, OCR Name: 33 HvH 405X405 SvS 272X272 SvH 643



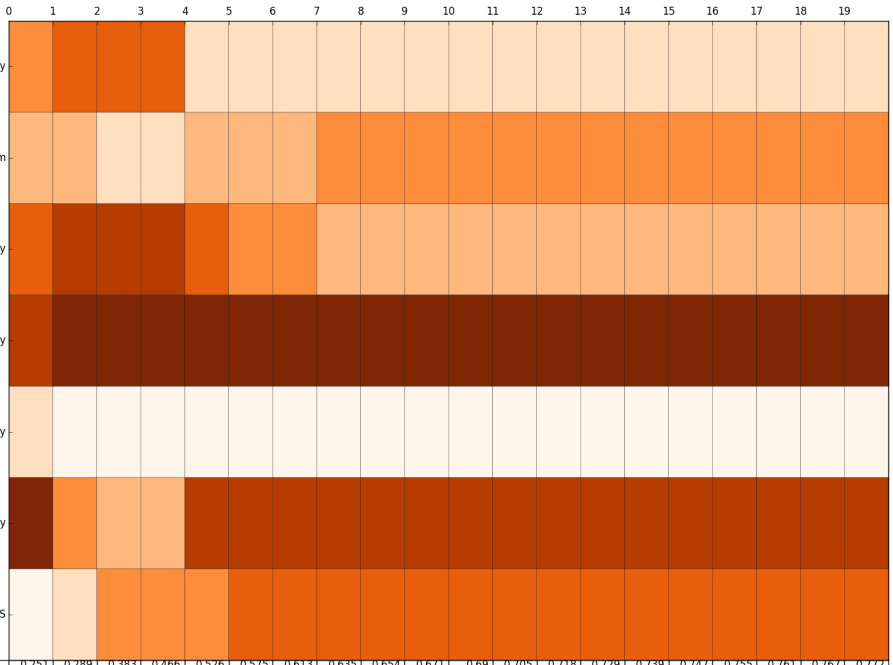
Soft Hybrid - Certainty	0.496	0.512	0.494	0.538	0.557	0.558	0.579	0.582	0.59	0.596	0.601	0.612	0.617	0.627	0.636	0.644	0.652	0.66	0.668	0.671
Random	0.559	0.579	0.594	0.596	0.608	0.6	0.607	0.612	0.62	0.622	0.616	0.621	0.625	0.629	0.635	0.639	0.643	0.647	0.647	0.652
MDS - Certainty	0.51	0.52	0.541	0.557	0.564	0.581	0.588	0.604	0.608	0.611	0.614	0.629	0.634	0.639	0.648	0.65	0.655	0.662	0.666	0.673
MDS - Uncertainty	0.473	0.53	0.531	0.549	0.546	0.535	0.551	0.562	0.572	0.57	0.576	0.579	0.589	0.593	0.6	0.601	0.603	0.608	0.616	0.62
Hard Hybrid - Certainty	0.537	0.585	0.58	0.598	0.588	0.592	0.606	0.612	0.619	0.621	0.628	0.635	0.641	0.645	0.652	0.655	0.658	0.663	0.668	0.674
Hard Hybrid - Uncertainty	0.554	0.529	0.541	0.567	0.564	0.583	0.593	0.606	0.612	0.614	0.622	0.625	0.625	0.634	0.635	0.635	0.636	0.638	0.646	0.651
BCS	0.588	0.547	0.546	0.57	0.572	0.578	0.592	0.603	0.615	0.621	0.629	0.637	0.638	0.642	0.647	0.648	0.647	0.65	0.653	0.657

BFMTV_BFMStory_2012-02-20_175800, Head Track: 871, Speaker Track: 342, OCR Name: 30 HvH 450X450 SvS 267X267 SvH 710



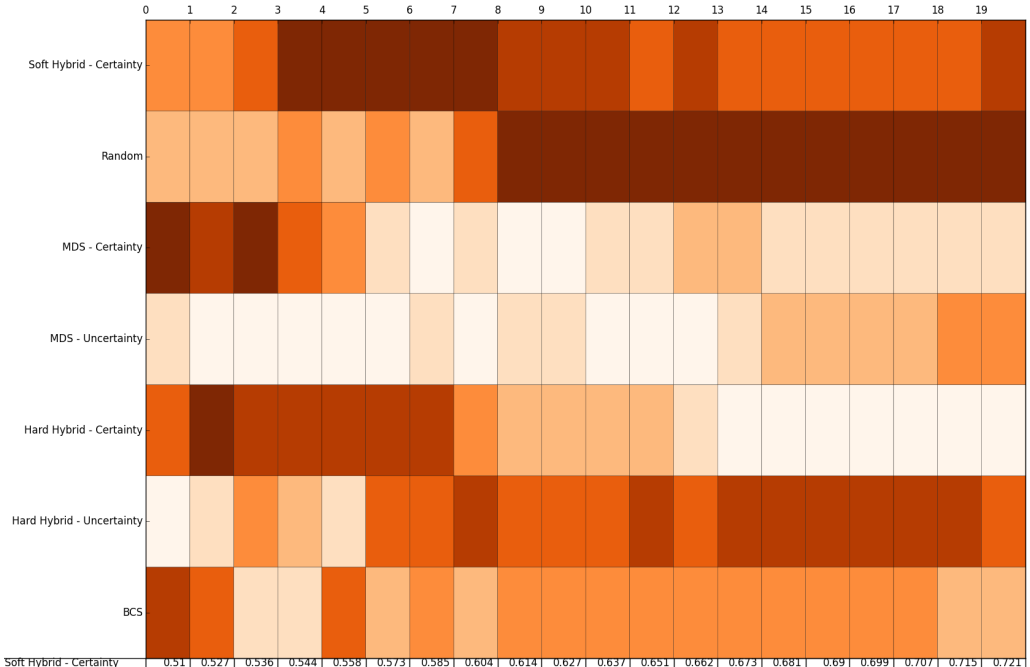
Soft Hybrid - Certainty	0.586	0.606	0.587	0.61	0.615	0.616	0.622	0.622	0.631	0.635	0.64	0.644	0.644	0.643	0.644	0.647	0.651	0.654	0.654	0.653
Random	0.604	0.615	0.63	0.633	0.63	0.62	0.626	0.633	0.635	0.642	0.641	0.644	0.651	0.655	0.658	0.661	0.664	0.666	0.672	0.678
MDS - Certainty	0.589	0.587	0.589	0.608	0.604	0.61	0.621	0.624	0.633	0.631	0.634	0.636	0.645	0.643	0.647	0.652	0.654	0.658	0.662	0.664
MDS - Uncertainty	0.562	0.589	0.605	0.604	0.621	0.63	0.632	0.638	0.641	0.643	0.645	0.652	0.65	0.651	0.655	0.659	0.662	0.666	0.668	0.673
Hard Hybrid - Certainty	0.601	0.582	0.609	0.589	0.6	0.611	0.613	0.61	0.614	0.623	0.626	0.631	0.637	0.641	0.641	0.646	0.649	0.651	0.655	0.658
Hard Hybrid - Uncertainty	0.58	0.617	0.629	0.62	0.621	0.625	0.628	0.632	0.637	0.633	0.633	0.64	0.647	0.653	0.653	0.662	0.668	0.669	0.671	0.672
BCS	0.632	0.64	0.641	0.643	0.65	0.652	0.652	0.652	0.659	0.661	0.66	0.669	0.673	0.679	0.68	0.68	0.685	0.682	0.68	0.682

BFMTV_CultureEtVous_2012-01-13_065040, Head Track: 69, Speaker Track: 38, OCR Name: 3 HvH 29X29 SvS 28X28 SvH 42



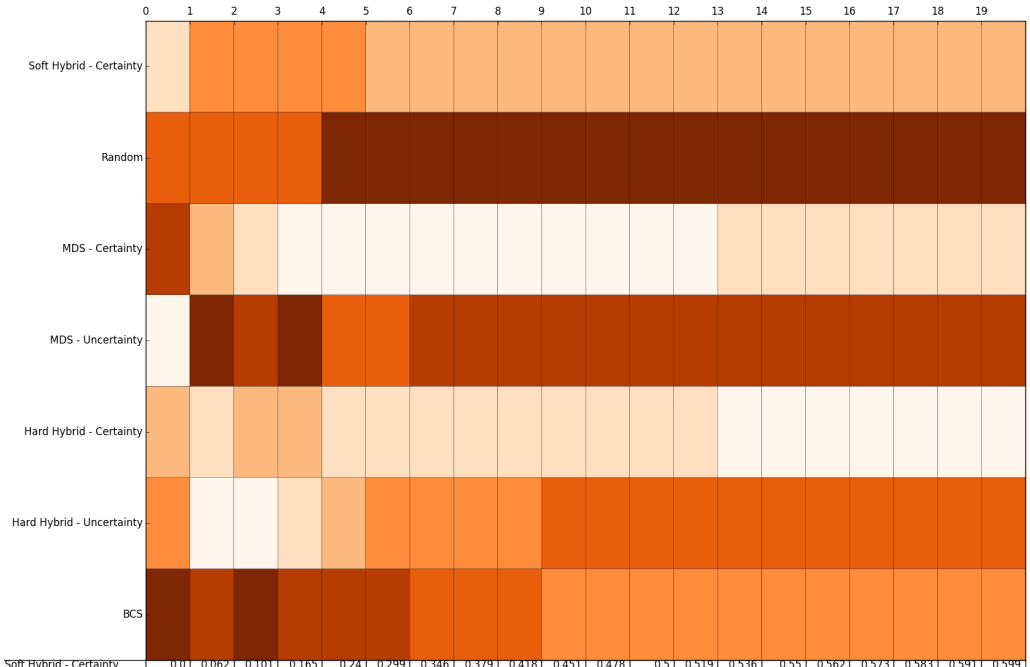
Soft Hybrid - Certainty	0.251	0.289	0.383	0.466	0.526	0.575	0.613	0.635	0.654	0.671	0.69	0.705	0.718	0.729	0.739	0.747	0.755	0.761	0.767	0.772
Random	0.256	0.394	0.455	0.504	0.524	0.559	0.584	0.605	0.624	0.647	0.664	0.679	0.693	0.705	0.715	0.724	0.731	0.738	0.744	0.75
MDS - Certainty	0.241	0.281	0.361	0.432	0.501	0.546	0.582	0.614	0.64	0.66	0.676	0.691	0.706	0.718	0.729	0.738	0.747	0.754	0.76	0.766
MDS - Uncertainty	0.235	0.271	0.311	0.362	0.426	0.453	0.486	0.504	0.53	0.553	0.576	0.596	0.618	0.637	0.653	0.667	0.679	0.69	0.7	0.709
Hard Hybrid - Certainty	0.267	0.507	0.592	0.63	0.655	0.659	0.672	0.683	0.694	0.705	0.718	0.731	0.742	0.752	0.759	0.766	0.773	0.778	0.784	0.788
Hard Hybrid - Uncertainty	0.23	0.38	0.453	0.476	0.488	0.507	0.521	0.529	0.551	0.573	0.592	0.614	0.635	0.653	0.667	0.679	0.691	0.701	0.71	0.718
BCS	0.276	0.397	0.448	0.467	0.505	0.519	0.532	0.543	0.568	0.594	0.617	0.637	0.656	0.671	0.685	0.697	0.708	0.717	0.726	0.733

BFMTV_CultureEtVous_2012-01-16_075000, Head Track: 163, Speaker Track: 42, OCR Name: 5 HvH 78X78 SvS 32X32 SvH 117



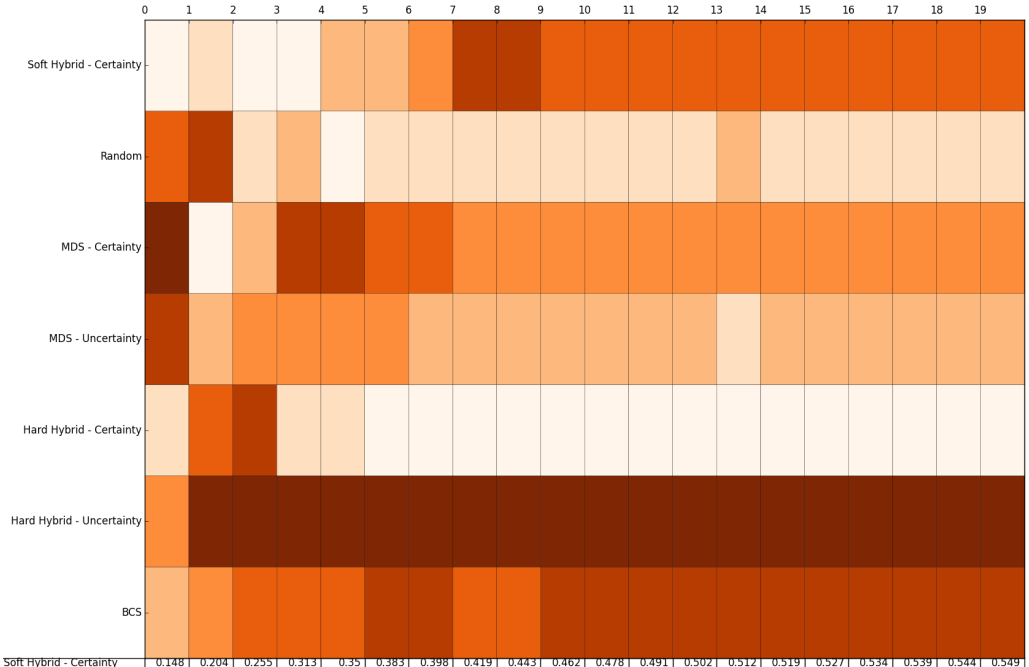
Soft Hybrid - Certainty	0.51	0.527	0.536	0.544	0.558	0.573	0.585	0.604	0.614	0.627	0.637	0.651	0.662	0.673	0.681	0.69	0.699	0.707	0.715	0.721
Random	0.512	0.54	0.55	0.558	0.576	0.584	0.6	0.609	0.612	0.624	0.635	0.644	0.655	0.663	0.668	0.677	0.683	0.69	0.698	0.704
MDS - Certainty	0.5	0.509	0.528	0.554	0.576	0.592	0.609	0.62	0.635	0.647	0.655	0.664	0.672	0.683	0.694	0.702	0.708	0.715	0.723	0.729
MDS - Uncertainty	0.525	0.542	0.553	0.565	0.58	0.597	0.608	0.621	0.634	0.646	0.655	0.665	0.675	0.683	0.691	0.699	0.707	0.713	0.718	0.724
Hard Hybrid - Certainty	0.508	0.503	0.53	0.548	0.562	0.578	0.594	0.609	0.624	0.635	0.65	0.662	0.675	0.687	0.696	0.705	0.714	0.722	0.73	0.736
Hard Hybrid - Uncertainty	0.537	0.541	0.549	0.561	0.577	0.583	0.596	0.605	0.618	0.63	0.639	0.651	0.663	0.671	0.68	0.689	0.697	0.706	0.714	0.722
BCS	0.502	0.519	0.551	0.564	0.574	0.585	0.597	0.611	0.622	0.633	0.645	0.657	0.667	0.677	0.686	0.697	0.706	0.713	0.72	0.727

BFMTV_CultureEtVous_2012-01-17_065040, Head Track: 48, Speaker Track: 49, OCR Name: 5 HvH 23X23 SvS 35X35 SvH 30



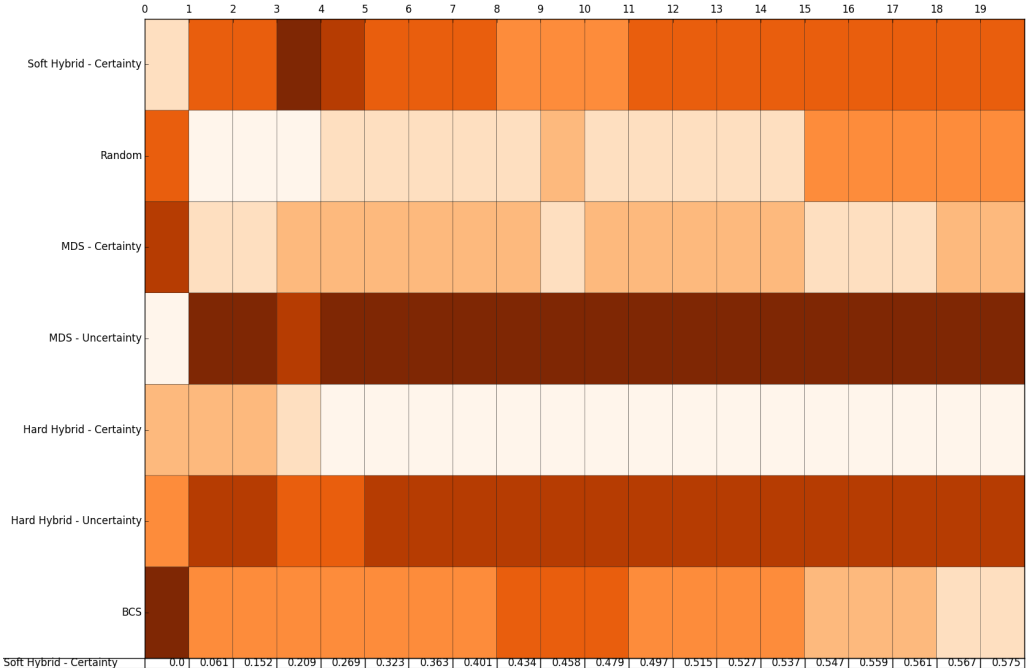
Soft Hybrid - Certainty	0.0	0.062	0.101	0.165	0.24	0.299	0.346	0.379	0.418	0.451	0.478	0.5	0.519	0.536	0.55	0.562	0.573	0.583	0.591	0.599
Random	0.0	0.047	0.075	0.117	0.146	0.174	0.22	0.257	0.297	0.333	0.367	0.398	0.419	0.443	0.462	0.48	0.495	0.506	0.518	0.527
MDS - Certainty	0.0	0.069	0.184	0.265	0.322	0.372	0.423	0.461	0.491	0.511	0.531	0.548	0.562	0.574	0.584	0.593	0.602	0.609	0.615	0.621
MDS - Uncertainty	0.0	0.026	0.072	0.095	0.161	0.221	0.246	0.278	0.315	0.35	0.383	0.412	0.437	0.458	0.476	0.492	0.506	0.518	0.53	0.54
Hard Hybrid - Certainty	0.0	0.094	0.146	0.233	0.305	0.355	0.404	0.445	0.478	0.506	0.528	0.546	0.561	0.575	0.586	0.596	0.605	0.613	0.62	0.626
Hard Hybrid - Uncertainty	0.0	0.164	0.26	0.249	0.242	0.284	0.308	0.336	0.371	0.406	0.435	0.459	0.479	0.497	0.512	0.525	0.537	0.547	0.557	0.565
BCS	0.0	0.036	0.07	0.115	0.147	0.212	0.264	0.323	0.371	0.409	0.438	0.464	0.486	0.504	0.521	0.535	0.547	0.558	0.568	0.577

BFMTV_CultureEtVous_2012-01-18_065040, Head Track: 102, Speaker Track: 47, OCR Name: 6 HvH 41X41 SvS 37X37 SvH 42



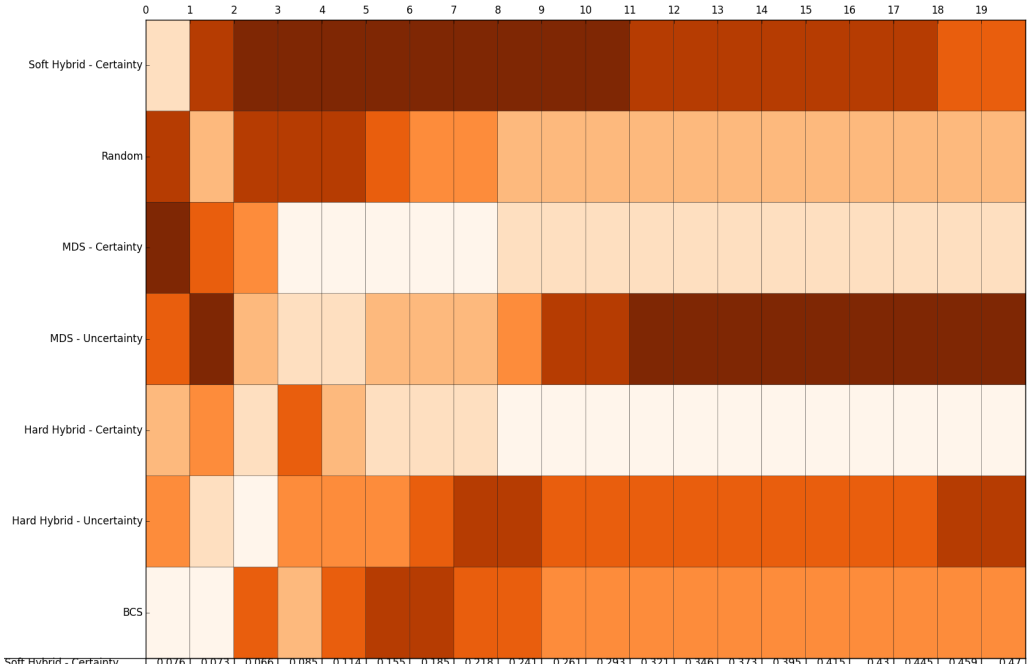
Soft Hybrid - Certainty	0.148	0.204	0.255	0.313	0.35	0.383	0.398	0.419	0.443	0.462	0.478	0.491	0.502	0.512	0.519	0.527	0.534	0.539	0.544	0.549
Random	0.136	0.17	0.249	0.309	0.366	0.395	0.421	0.447	0.47	0.487	0.499	0.512	0.519	0.527	0.537	0.544	0.551	0.556	0.561	0.565
MDS - Certainty	0.126	0.213	0.238	0.273	0.31	0.36	0.397	0.429	0.453	0.471	0.483	0.496	0.508	0.518	0.526	0.535	0.542	0.548	0.553	0.558
MDS - Uncertainty	0.127	0.196	0.234	0.299	0.347	0.374	0.413	0.442	0.465	0.48	0.496	0.507	0.518	0.528	0.536	0.542	0.549	0.555	0.56	0.564
Hard Hybrid - Certainty	0.147	0.177	0.224	0.313	0.362	0.399	0.432	0.455	0.477	0.495	0.508	0.518	0.526	0.534	0.54	0.546	0.551	0.556	0.561	0.565
Hard Hybrid - Uncertainty	0.141	0.163	0.207	0.24	0.295	0.343	0.378	0.404	0.427	0.447	0.466	0.481	0.493	0.504	0.513	0.521	0.529	0.536	0.542	0.547
BCS	0.145	0.178	0.225	0.289	0.333	0.356	0.396	0.422	0.444	0.461	0.476	0.488	0.498	0.508	0.518	0.525	0.531	0.537	0.543	0.548

BFMTV_CultureEtVous_2012-01-19_065040, Head Track: 78, Speaker Track: 41, OCR Name: 4 HvH 43X43 SvS 31X31 SvH 52



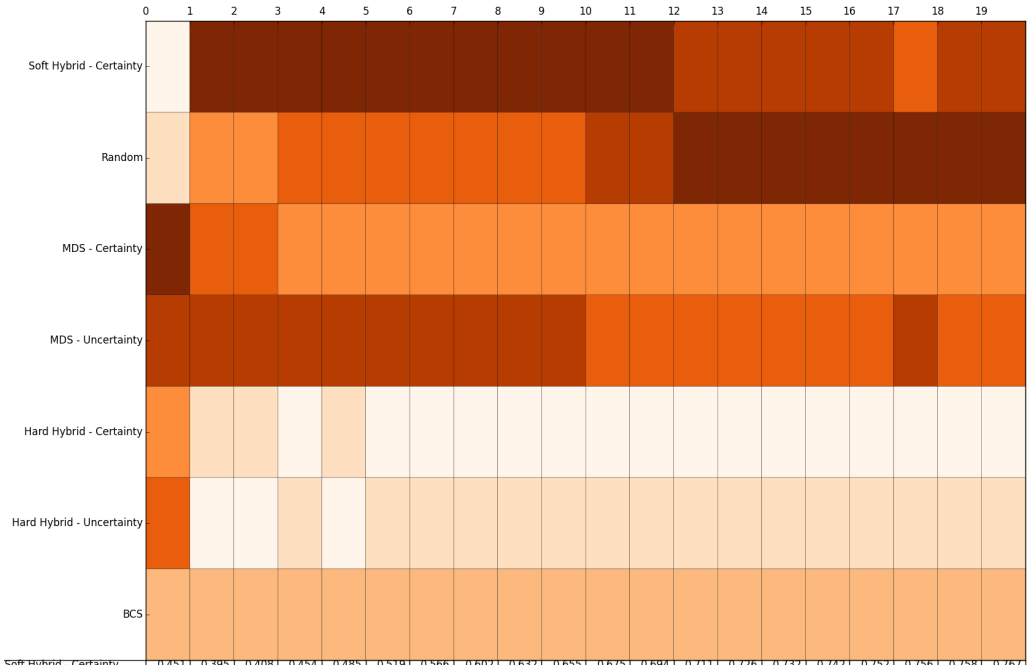
Soft Hybrid - Certainty	0.0	0.061	0.152	0.209	0.269	0.323	0.363	0.401	0.434	0.458	0.479	0.497	0.515	0.527	0.537	0.547	0.559	0.561	0.567	0.575
Random	0.0	0.189	0.267	0.31	0.346	0.384	0.416	0.447	0.465	0.477	0.497	0.511	0.524	0.538	0.548	0.556	0.565	0.57	0.576	0.58
MDS - Certainty	0.0	0.123	0.187	0.273	0.314	0.365	0.405	0.435	0.457	0.478	0.494	0.511	0.524	0.537	0.548	0.559	0.569	0.578	0.581	0.587
MDS - Uncertainty	0.0	0.04	0.144	0.211	0.239	0.275	0.302	0.34	0.37	0.402	0.429	0.454	0.475	0.49	0.505	0.52	0.532	0.54	0.546	0.554
Hard Hybrid - Certainty	0.0	0.078	0.185	0.287	0.351	0.396	0.421	0.448	0.471	0.486	0.506	0.523	0.54	0.553	0.563	0.569	0.578	0.586	0.594	0.601
Hard Hybrid - Uncertainty	0.0	0.045	0.152	0.232	0.277	0.321	0.343	0.376	0.409	0.433	0.451	0.469	0.487	0.504	0.515	0.53	0.54	0.553	0.561	0.572
BCS	0.0	0.076	0.178	0.25	0.31	0.345	0.366	0.402	0.424	0.455	0.475	0.5	0.517	0.532	0.546	0.559	0.569	0.578	0.583	0.59

BFMTV_CultureEtVous_2012-02-14_065040, Head Track: 104, Speaker Track: 48, OCR Name: 9 HvH 48X48 SvS 39X39 SvH 61



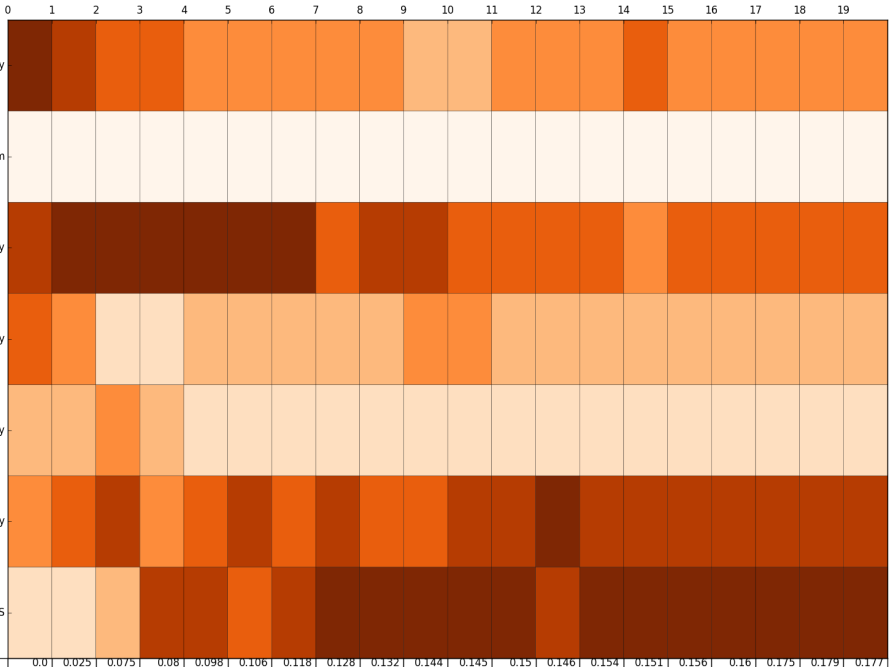
Soft Hybrid - Certainty	0.076	0.073	0.066	0.085	0.114	0.155	0.185	0.218	0.241	0.261	0.293	0.321	0.346	0.373	0.395	0.415	0.43	0.445	0.459	0.47
Random	0.025	0.083	0.079	0.088	0.151	0.181	0.21	0.234	0.268	0.302	0.326	0.353	0.38	0.4	0.419	0.434	0.445	0.458	0.469	0.48
MDS - Certainty	0.025	0.074	0.112	0.176	0.22	0.258	0.285	0.306	0.328	0.355	0.37	0.393	0.411	0.431	0.448	0.461	0.472	0.483	0.491	0.5
MDS - Uncertainty	0.029	0.07	0.114	0.151	0.186	0.191	0.215	0.243	0.265	0.269	0.295	0.318	0.338	0.362	0.381	0.4	0.416	0.43	0.442	0.453
Hard Hybrid - Certainty	0.071	0.08	0.118	0.134	0.176	0.223	0.261	0.3	0.34	0.36	0.382	0.408	0.429	0.45	0.467	0.482	0.496	0.508	0.517	0.527
Hard Hybrid - Uncertainty	0.064	0.114	0.127	0.135	0.169	0.189	0.204	0.233	0.251	0.272	0.302	0.329	0.355	0.378	0.397	0.417	0.434	0.447	0.458	0.47
BCS	0.084	0.128	0.111	0.141	0.153	0.177	0.195	0.234	0.254	0.281	0.315	0.34	0.367	0.389	0.409	0.427	0.443	0.455	0.468	0.479

BFMTV_CultureEtVous_2012-02-15_065040, Head Track: 180, Speaker Track: 34, OCR Name: 4 HvH 67X67 SvS 27X27 SvH 113



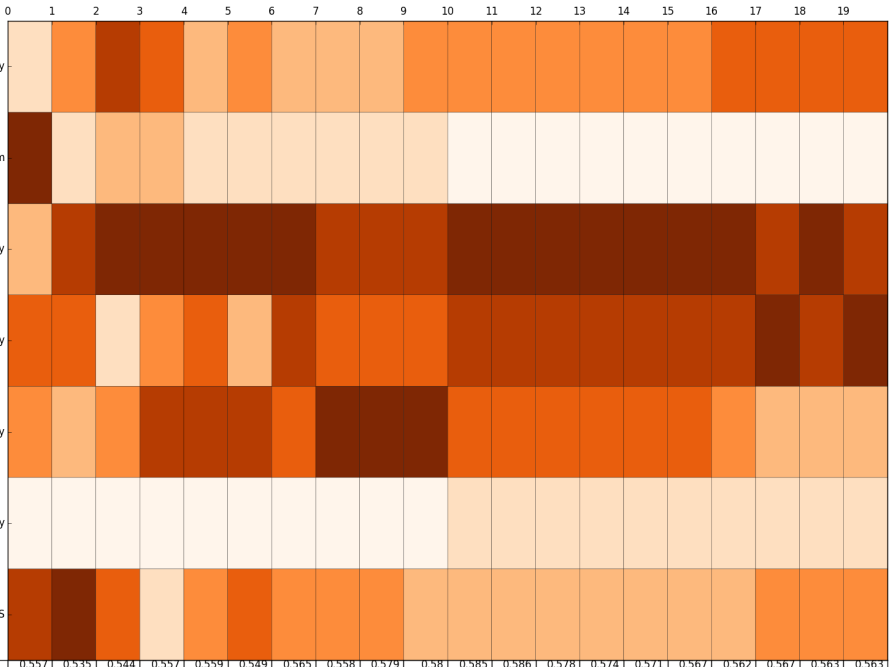
Soft Hybrid - Certainty	0.451	0.395	0.408	0.454	0.485	0.519	0.566	0.602	0.632	0.655	0.675	0.694	0.711	0.726	0.732	0.742	0.752	0.756	0.758	0.767
Random	0.447	0.479	0.519	0.539	0.563	0.589	0.622	0.651	0.668	0.676	0.69	0.7	0.706	0.716	0.719	0.729	0.734	0.744	0.748	0.751
MDS - Certainty	0.35	0.423	0.491	0.547	0.587	0.631	0.659	0.679	0.699	0.716	0.727	0.74	0.746	0.759	0.766	0.775	0.775	0.779	0.782	0.786
MDS - Uncertainty	0.35	0.415	0.465	0.508	0.536	0.571	0.605	0.631	0.653	0.672	0.691	0.71	0.721	0.732	0.741	0.749	0.754	0.755	0.765	0.773
Hard Hybrid - Certainty	0.414	0.499	0.544	0.612	0.656	0.704	0.728	0.745	0.759	0.773	0.784	0.79	0.795	0.804	0.81	0.81	0.813	0.812	0.815	0.82
Hard Hybrid - Uncertainty	0.397	0.515	0.578	0.605	0.657	0.683	0.709	0.731	0.741	0.756	0.759	0.769	0.777	0.781	0.783	0.788	0.789	0.795	0.802	0.809
BCS	0.431	0.494	0.534	0.569	0.611	0.65	0.674	0.694	0.715	0.726	0.742	0.753	0.765	0.771	0.776	0.782	0.787	0.791	0.795	0.798

LCP_CaVousRegarde_2011-12-20_204600, Head Track: 397, Speaker Track: 332, OCR Name: 30 HvH 282X282 SvS 268X268 SvH 406



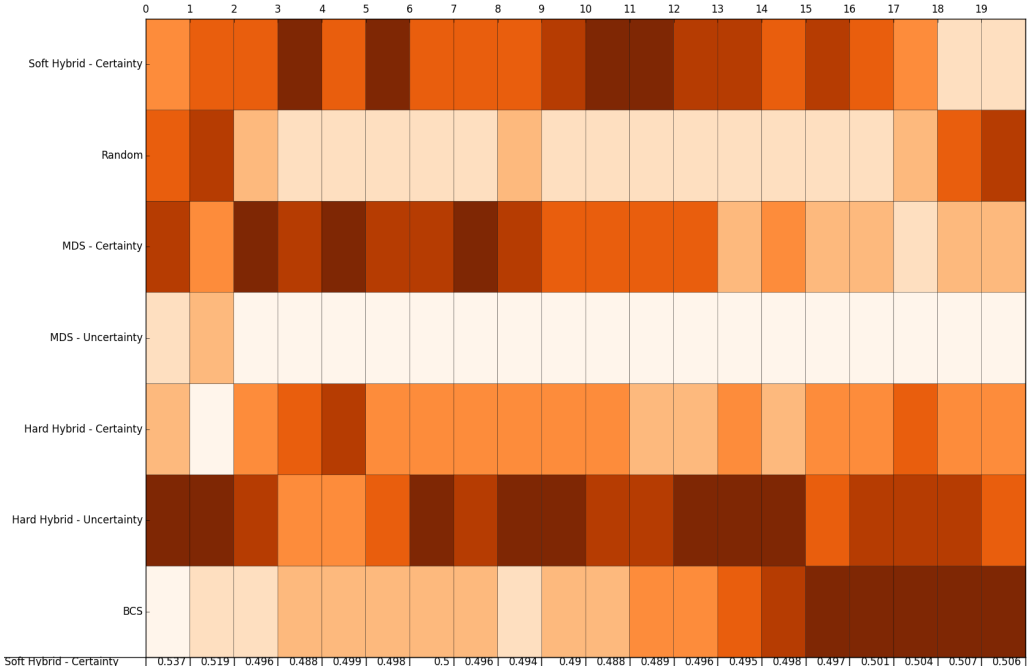
Soft Hybrid - Certainty	0.0	0.025	0.075	0.08	0.098	0.106	0.118	0.128	0.132	0.144	0.145	0.15	0.146	0.154	0.151	0.156	0.16	0.175	0.179	0.177
Random	0.116	0.119	0.197	0.186	0.203	0.235	0.256	0.245	0.25	0.257	0.256	0.264	0.26	0.268	0.287	0.295	0.299	0.307	0.309	0.316
MDS - Certainty	0.007	0.011	0.042	0.066	0.065	0.09	0.097	0.118	0.115	0.121	0.122	0.132	0.142	0.142	0.152	0.156	0.152	0.158	0.16	0.166
MDS - Uncertainty	0.057	0.087	0.116	0.095	0.121	0.121	0.135	0.13	0.136	0.136	0.141	0.15	0.148	0.159	0.162	0.164	0.174	0.182	0.187	0.186
Hard Hybrid - Certainty	0.074	0.09	0.088	0.086	0.122	0.164	0.175	0.172	0.18	0.176	0.176	0.173	0.185	0.198	0.195	0.192	0.201	0.208	0.215	0.214
Hard Hybrid - Uncertainty	0.068	0.055	0.063	0.081	0.081	0.095	0.112	0.104	0.122	0.122	0.119	0.126	0.122	0.126	0.125	0.13	0.136	0.14	0.145	0.152
BCS	0.11	0.103	0.107	0.08	0.081	0.098	0.101	0.102	0.1	0.113	0.115	0.124	0.128	0.123	0.119	0.124	0.129	0.132	0.145	0.149

LCP_CaVousRegarde_2012-01-19_204700, Head Track: 371, Speaker Track: 290, OCR Name: 22 HvH 244X244 SvS 229X229 SvH 384



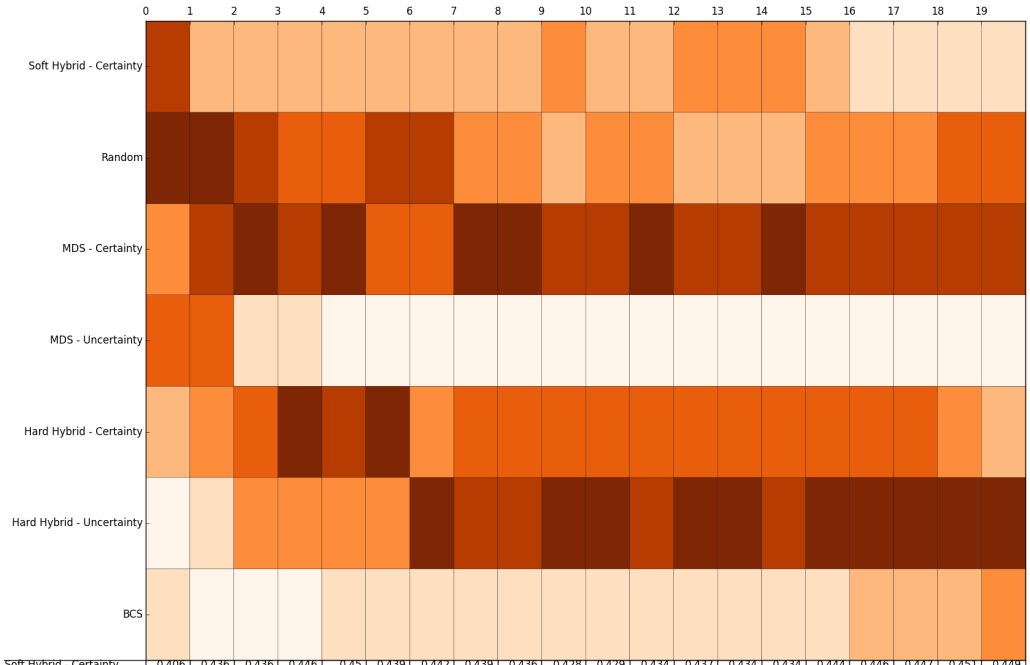
Soft Hybrid - Certainty	0.557	0.535	0.544	0.557	0.559	0.549	0.565	0.558	0.579	0.58	0.585	0.586	0.578	0.574	0.571	0.567	0.562	0.567	0.563	0.563
Random	0.496	0.563	0.568	0.567	0.586	0.597	0.609	0.607	0.607	0.605	0.603	0.612	0.617	0.628	0.632	0.637	0.639	0.637	0.637	0.635
MDS - Certainty	0.552	0.529	0.495	0.511	0.501	0.522	0.537	0.544	0.55	0.553	0.551	0.55	0.546	0.55	0.551	0.543	0.537	0.541	0.537	0.545
MDS - Uncertainty	0.537	0.534	0.568	0.561	0.545	0.55	0.537	0.548	0.562	0.561	0.555	0.554	0.552	0.552	0.555	0.547	0.541	0.54	0.538	0.535
Hard Hybrid - Certainty	0.547	0.554	0.562	0.554	0.543	0.527	0.542	0.543	0.549	0.552	0.559	0.556	0.558	0.559	0.557	0.558	0.564	0.571	0.575	0.571
Hard Hybrid - Uncertainty	0.641	0.657	0.636	0.645	0.634	0.628	0.626	0.615	0.613	0.605	0.599	0.592	0.591	0.59	0.589	0.587	0.589	0.587	0.581	0.582
BCS	0.516	0.507	0.555	0.572	0.554	0.535	0.544	0.549	0.565	0.587	0.592	0.588	0.584	0.58	0.575	0.568	0.573	0.57	0.568	0.567

LCP_CaVousRegarde_2012-01-25_204700, Head Track: 428, Speaker Track: 352, OCR Name: 25 HvH 281X281 SvS 282X282 SvH 429



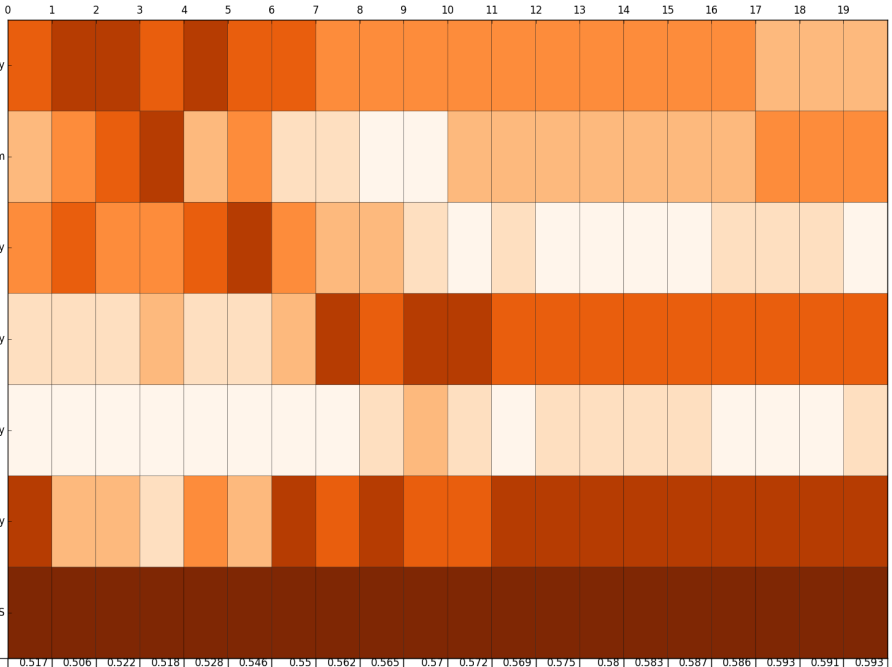
Soft Hybrid - Certainty	0.537	0.519	0.496	0.488	0.499	0.498	0.5	0.496	0.494	0.49	0.488	0.489	0.496	0.495	0.498	0.497	0.501	0.504	0.507	0.506
Random	0.531	0.498	0.539	0.549	0.549	0.547	0.53	0.526	0.524	0.523	0.521	0.524	0.523	0.52	0.514	0.511	0.509	0.506	0.501	0.5
MDS - Certainty	0.53	0.54	0.485	0.49	0.487	0.5	0.5	0.494	0.492	0.494	0.503	0.502	0.502	0.508	0.508	0.509	0.508	0.509	0.506	0.505
MDS - Uncertainty	0.585	0.552	0.561	0.565	0.554	0.552	0.551	0.551	0.548	0.542	0.541	0.527	0.529	0.53	0.531	0.535	0.538	0.536	0.533	0.531
Hard Hybrid - Certainty	0.559	0.56	0.511	0.496	0.498	0.506	0.516	0.513	0.507	0.51	0.509	0.513	0.509	0.507	0.508	0.506	0.503	0.503	0.503	0.503
Hard Hybrid - Uncertainty	0.425	0.482	0.487	0.51	0.514	0.505	0.493	0.496	0.489	0.49	0.494	0.493	0.49	0.494	0.494	0.499	0.501	0.5	0.499	0.5
BCS	0.605	0.556	0.548	0.536	0.528	0.526	0.519	0.522	0.525	0.521	0.518	0.513	0.504	0.498	0.497	0.497	0.49	0.49	0.485	0.487

LCP_EntreLesLignes_2011-12-16_192700, Head Track: 560, Speaker Track: 253, OCR Name: 15 HvH 345X345 SvS 203X203 SvH 493



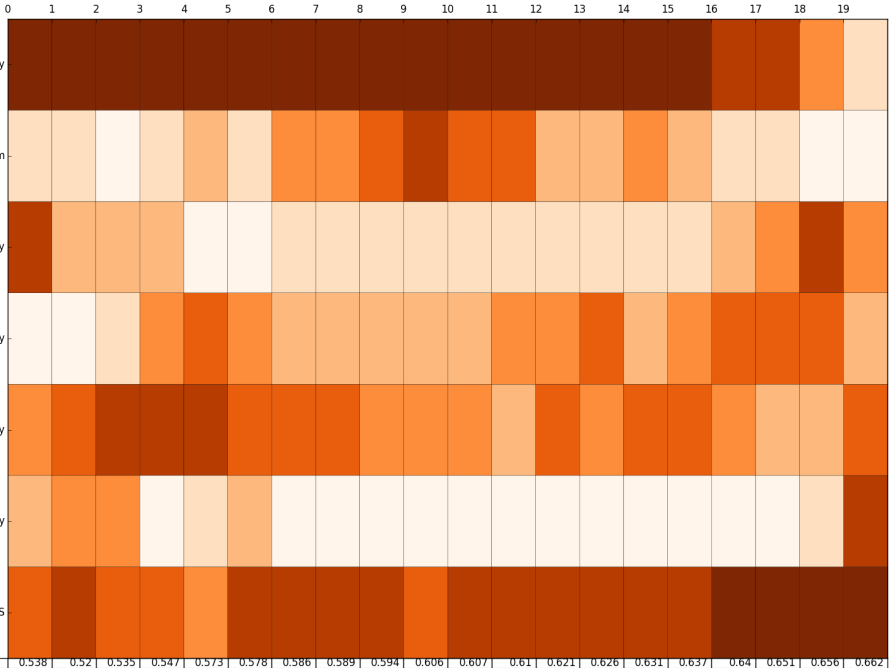
Soft Hybrid - Certainty	0.406	0.436	0.436	0.446	0.45	0.439	0.442	0.439	0.436	0.428	0.429	0.434	0.437	0.434	0.434	0.444	0.446	0.447	0.451	0.449
Random	0.327	0.385	0.402	0.404	0.406	0.404	0.402	0.415	0.428	0.43	0.428	0.43	0.439	0.444	0.444	0.443	0.443	0.442	0.438	0.44
MDS - Certainty	0.44	0.421	0.389	0.401	0.402	0.411	0.406	0.402	0.4	0.4	0.402	0.402	0.406	0.407	0.406	0.408	0.415	0.422	0.418	0.418
MDS - Uncertainty	0.407	0.422	0.437	0.457	0.472	0.468	0.465	0.463	0.458	0.461	0.454	0.452	0.454	0.458	0.46	0.457	0.463	0.461	0.458	0.458
Hard Hybrid - Certainty	0.459	0.427	0.414	0.4	0.402	0.403	0.408	0.405	0.415	0.413	0.419	0.425	0.432	0.433	0.434	0.435	0.437	0.434	0.44	0.445
Hard Hybrid - Uncertainty	0.493	0.448	0.423	0.409	0.406	0.413	0.398	0.405	0.402	0.399	0.4	0.402	0.402	0.404	0.407	0.403	0.406	0.41	0.409	0.412
BCS	0.47	0.481	0.479	0.47	0.461	0.448	0.443	0.442	0.444	0.443	0.445	0.447	0.443	0.445	0.449	0.452	0.446	0.447	0.445	0.444

LCP_EntreLesLignes_2012-01-27_192700, Head Track: 599, Speaker Track: 267, OCR Name: 14 HvH 329X329 SvS 210X210 SvH 460



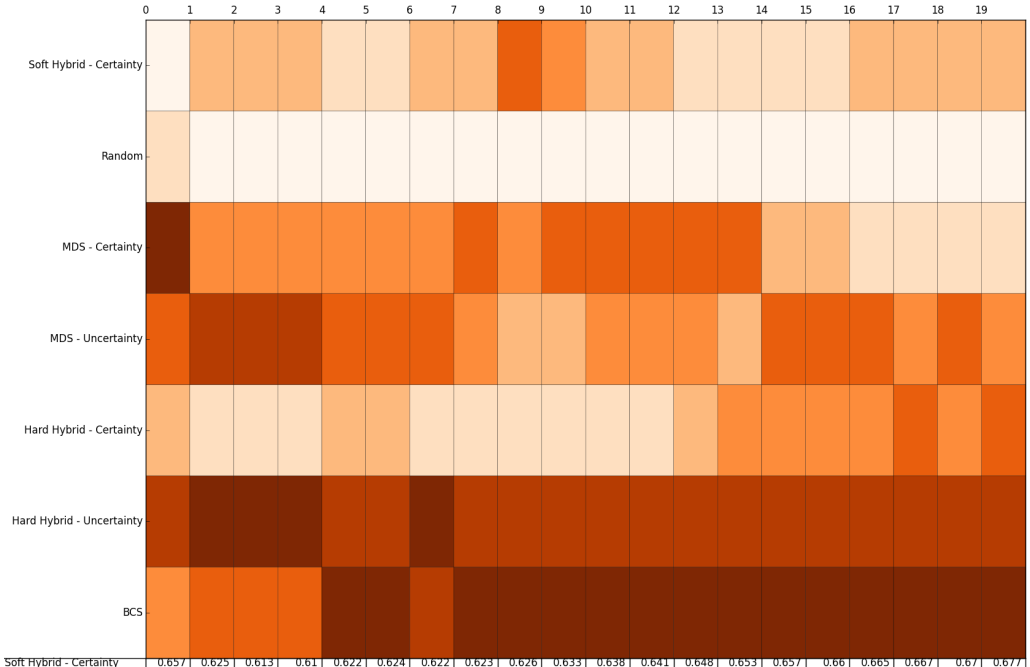
Soft Hybrid - Certainty	0.517	0.506	0.522	0.518	0.528	0.546	0.55	0.562	0.565	0.57	0.572	0.569	0.575	0.58	0.583	0.587	0.586	0.593	0.591	0.593
Random	0.57	0.554	0.537	0.515	0.545	0.55	0.572	0.571	0.575	0.584	0.579	0.575	0.583	0.589	0.589	0.59	0.59	0.584	0.585	0.583
MDS - Certainty	0.521	0.537	0.54	0.521	0.53	0.532	0.552	0.567	0.57	0.579	0.587	0.587	0.593	0.601	0.603	0.606	0.604	0.605	0.606	0.612
MDS - Uncertainty	0.587	0.597	0.566	0.551	0.551	0.564	0.559	0.555	0.556	0.554	0.551	0.56	0.57	0.567	0.563	0.567	0.574	0.577	0.58	0.578
Hard Hybrid - Certainty	0.65	0.608	0.592	0.59	0.588	0.577	0.579	0.578	0.572	0.578	0.584	0.591	0.589	0.595	0.597	0.6	0.604	0.609	0.606	0.611
Hard Hybrid - Uncertainty	0.498	0.587	0.558	0.568	0.542	0.555	0.543	0.557	0.556	0.558	0.555	0.557	0.557	0.555	0.554	0.554	0.558	0.559	0.563	0.563
BCS	0.489	0.488	0.506	0.504	0.517	0.508	0.506	0.513	0.517	0.513	0.513	0.518	0.516	0.521	0.528	0.533	0.541	0.54	0.547	0.546

LCP_EntreLesLignes_2012-05-11_192200, Head Track: 690, Speaker Track: 305, OCR Name: 13 HvH 310X310 SvS 240X240 SvH 596



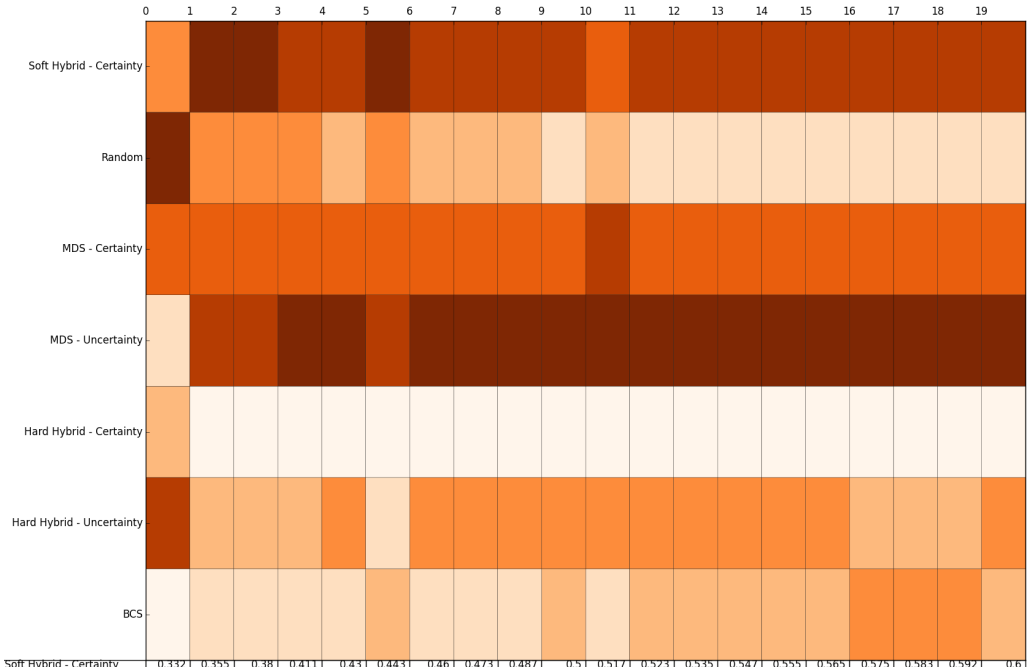
Soft Hybrid - Certainty	0.538	0.52	0.535	0.547	0.573	0.578	0.586	0.589	0.594	0.606	0.607	0.61	0.621	0.626	0.631	0.637	0.64	0.651	0.656	0.662
Random	0.619	0.622	0.64	0.617	0.622	0.63	0.61	0.613	0.619	0.614	0.63	0.64	0.644	0.648	0.646	0.654	0.656	0.656	0.661	0.663
MDS - Certainty	0.559	0.6	0.614	0.615	0.641	0.635	0.639	0.636	0.641	0.653	0.654	0.655	0.654	0.659	0.657	0.66	0.656	0.653	0.653	0.657
MDS - Uncertainty	0.653	0.649	0.631	0.614	0.607	0.616	0.625	0.632	0.629	0.638	0.642	0.64	0.643	0.641	0.648	0.652	0.653	0.651	0.655	0.658
Hard Hybrid - Certainty	0.572	0.561	0.558	0.586	0.591	0.603	0.598	0.607	0.624	0.627	0.636	0.641	0.636	0.642	0.645	0.647	0.655	0.655	0.658	0.655
Hard Hybrid - Uncertainty	0.575	0.592	0.607	0.623	0.628	0.626	0.643	0.654	0.661	0.665	0.666	0.668	0.658	0.659	0.658	0.663	0.667	0.668	0.66	0.653
BCS	0.56	0.542	0.597	0.604	0.612	0.583	0.594	0.603	0.605	0.615	0.615	0.625	0.626	0.629	0.639	0.638	0.639	0.644	0.648	0.652

LCP_LCPInfo13h30_2012-01-24_132700, Head Track: 212, Speaker Track: 163, OCR Name: 13 HvH 128X128 SvS 130X130 SvH 179



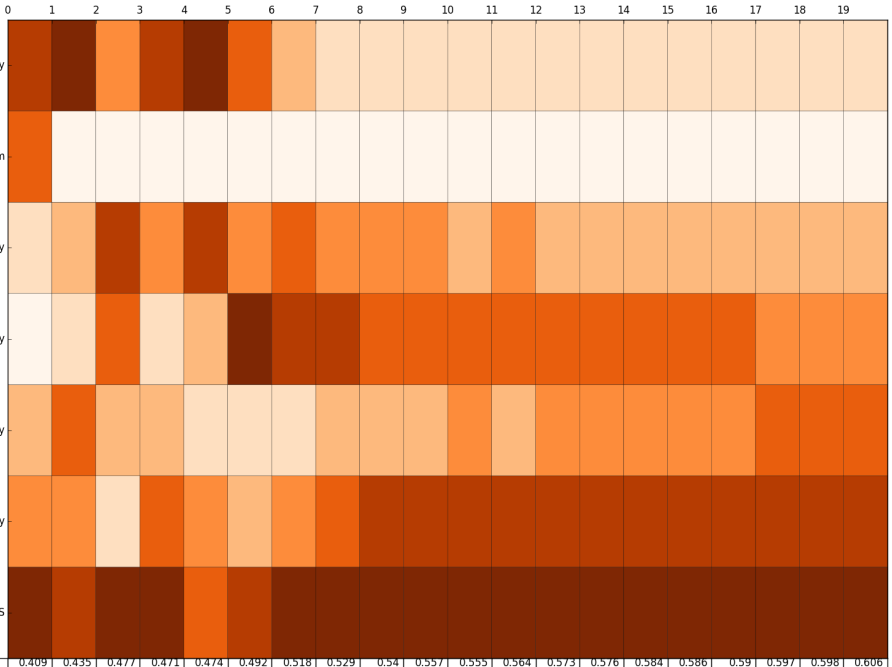
Soft Hybrid - Certainty	0.657	0.625	0.613	0.61	0.622	0.624	0.622	0.623	0.626	0.633	0.638	0.641	0.648	0.653	0.657	0.66	0.665	0.667	0.67	0.677
Random	0.652	0.639	0.635	0.625	0.637	0.644	0.636	0.641	0.647	0.646	0.653	0.659	0.663	0.664	0.664	0.664	0.667	0.673	0.677	0.682
MDS - Certainty	0.586	0.605	0.597	0.597	0.616	0.616	0.619	0.619	0.627	0.629	0.631	0.637	0.643	0.648	0.656	0.66	0.666	0.671	0.676	0.678
MDS - Uncertainty	0.605	0.592	0.578	0.577	0.589	0.595	0.602	0.62	0.627	0.634	0.637	0.639	0.645	0.65	0.649	0.652	0.653	0.657	0.659	0.663
Hard Hybrid - Certainty	0.63	0.63	0.625	0.615	0.621	0.623	0.628	0.635	0.641	0.642	0.643	0.645	0.647	0.649	0.652	0.654	0.655	0.657	0.66	0.661
Hard Hybrid - Uncertainty	0.6	0.533	0.551	0.567	0.577	0.585	0.585	0.593	0.599	0.602	0.611	0.62	0.626	0.631	0.639	0.644	0.65	0.652	0.656	0.661
BCS	0.606	0.599	0.584	0.578	0.577	0.585	0.59	0.589	0.595	0.597	0.6	0.599	0.602	0.604	0.603	0.607	0.612	0.613	0.615	0.618

LCP_LCPInfo13h30_2012-01-25_132700, Head Track: 331, Speaker Track: 172, OCR Name: 10 HvH 208X208 SvS 145X145 SvH 285



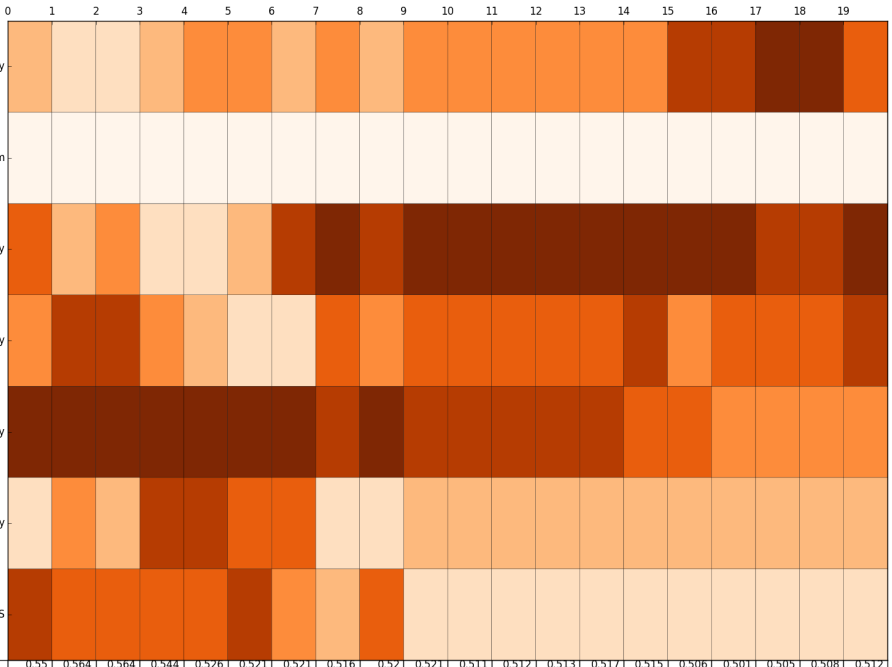
Soft Hybrid - Certainty	0.332	0.355	0.38	0.411	0.43	0.443	0.46	0.473	0.487	0.5	0.517	0.523	0.535	0.547	0.555	0.565	0.575	0.583	0.592	0.6
Random	0.323	0.375	0.409	0.435	0.463	0.483	0.502	0.525	0.539	0.557	0.564	0.579	0.594	0.605	0.617	0.629	0.636	0.643	0.649	0.653
MDS - Certainty	0.328	0.372	0.397	0.426	0.439	0.454	0.471	0.483	0.492	0.506	0.516	0.528	0.537	0.552	0.563	0.572	0.58	0.59	0.598	0.607
MDS - Uncertainty	0.342	0.364	0.386	0.405	0.424	0.443	0.456	0.468	0.476	0.484	0.486	0.496	0.506	0.515	0.522	0.531	0.538	0.547	0.556	0.566
Hard Hybrid - Certainty	0.336	0.414	0.469	0.516	0.539	0.556	0.574	0.587	0.596	0.607	0.614	0.623	0.632	0.637	0.645	0.652	0.658	0.665	0.671	0.676
Hard Hybrid - Uncertainty	0.326	0.377	0.418	0.437	0.461	0.491	0.498	0.512	0.531	0.539	0.551	0.558	0.57	0.58	0.586	0.594	0.6	0.606	0.612	0.617
BCS	0.366	0.408	0.423	0.451	0.476	0.491	0.517	0.536	0.548	0.554	0.567	0.573	0.579	0.586	0.592	0.596	0.599	0.605	0.611	0.619

LCP_LCPInfo13h30_2012-01-27_132700, Head Track: 306, Speaker Track: 177, OCR Name: 11 HvH 182X182 SvS 135X135 SvH 270



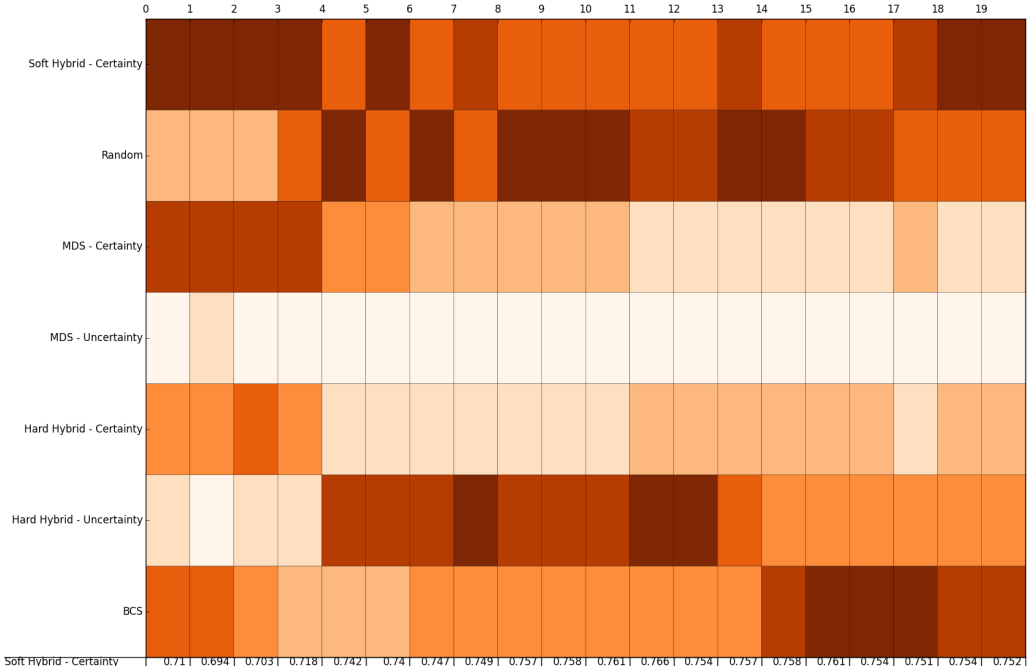
Soft Hybrid - Certainty	0.409	0.435	0.477	0.471	0.474	0.492	0.518	0.529	0.54	0.557	0.555	0.564	0.573	0.576	0.584	0.586	0.59	0.597	0.598	0.606
Random	0.458	0.509	0.545	0.543	0.569	0.586	0.594	0.595	0.599	0.603	0.61	0.618	0.62	0.625	0.624	0.624	0.63	0.628	0.626	0.627
MDS - Certainty	0.49	0.472	0.461	0.482	0.481	0.5	0.514	0.523	0.524	0.533	0.536	0.537	0.543	0.552	0.558	0.564	0.57	0.575	0.577	0.581
MDS - Uncertainty	0.512	0.48	0.474	0.505	0.494	0.488	0.495	0.509	0.516	0.514	0.517	0.524	0.528	0.536	0.539	0.545	0.552	0.559	0.564	0.568
Hard Hybrid - Certainty	0.472	0.47	0.481	0.49	0.499	0.52	0.524	0.527	0.53	0.535	0.535	0.539	0.54	0.548	0.553	0.554	0.553	0.555	0.559	0.564
Hard Hybrid - Uncertainty	0.469	0.471	0.484	0.475	0.493	0.511	0.517	0.512	0.504	0.502	0.488	0.49	0.49	0.501	0.503	0.513	0.516	0.521	0.527	0.53
BCS	0.408	0.467	0.442	0.466	0.49	0.491	0.49	0.473	0.47	0.463	0.467	0.463	0.459	0.453	0.454	0.453	0.456	0.462	0.471	0.479

LCP_PileEtFace_2011-11-19_085700, Head Track: 243, Speaker Track: 211, OCR Name: 40 HvH 173X173 SvS 162X162 SvH 289



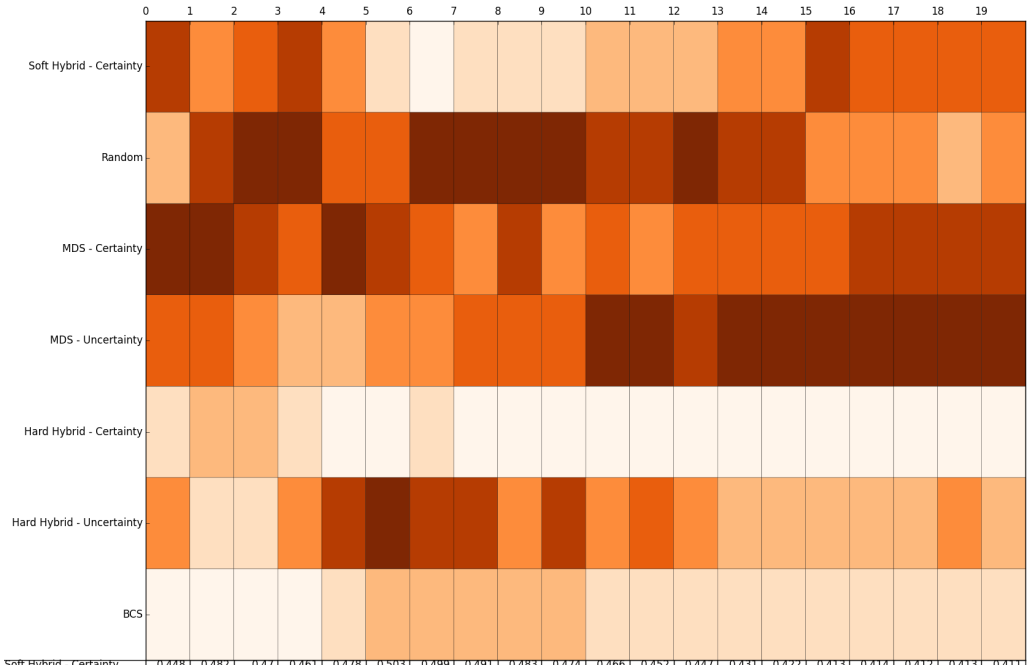
Soft Hybrid - Certainty	0.55	0.564	0.564	0.544	0.526	0.521	0.521	0.516	0.52	0.521	0.511	0.512	0.513	0.517	0.515	0.506	0.501	0.505	0.508	0.512
Random	0.591	0.567	0.57	0.558	0.589	0.589	0.598	0.598	0.585	0.586	0.594	0.589	0.587	0.579	0.58	0.575	0.566	0.571	0.574	0.57
MDS - Certainty	0.459	0.553	0.538	0.545	0.547	0.523	0.493	0.494	0.486	0.477	0.485	0.488	0.498	0.499	0.499	0.5	0.501	0.507	0.509	0.511
MDS - Uncertainty	0.532	0.485	0.462	0.532	0.538	0.527	0.533	0.515	0.513	0.511	0.506	0.503	0.509	0.514	0.509	0.511	0.512	0.51	0.512	0.511
Hard Hybrid - Certainty	0.376	0.466	0.441	0.447	0.471	0.471	0.482	0.495	0.48	0.487	0.487	0.49	0.5	0.504	0.512	0.511	0.522	0.523	0.526	0.523
Hard Hybrid - Uncertainty	0.576	0.53	0.548	0.496	0.499	0.519	0.5	0.526	0.531	0.523	0.526	0.522	0.52	0.52	0.523	0.532	0.53	0.533	0.538	0.537
BCS	0.456	0.522	0.534	0.497	0.503	0.515	0.516	0.517	0.513	0.538	0.532	0.543	0.548	0.534	0.542	0.543	0.548	0.544	0.542	0.539

LCP_PileEtFace_2011-12-01_192700, Head Track: 389, Speaker Track: 255, OCR Name: 51 HvH 250X250 SvS 197X197 SvH 349



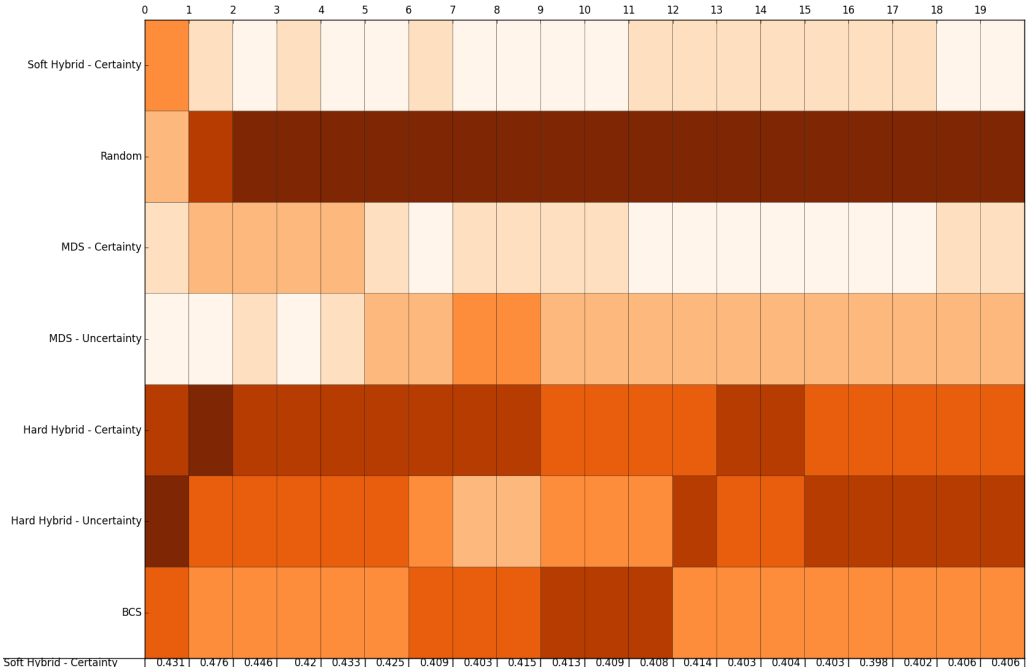
Soft Hybrid - Certainty	0.71	0.694	0.703	0.718	0.742	0.74	0.747	0.749	0.757	0.758	0.761	0.766	0.754	0.757	0.758	0.761	0.754	0.751	0.754	0.752
Random	0.796	0.784	0.784	0.764	0.734	0.748	0.741	0.75	0.741	0.74	0.738	0.747	0.752	0.756	0.749	0.754	0.753	0.758	0.759	0.758
MDS - Certainty	0.724	0.72	0.732	0.75	0.745	0.755	0.768	0.77	0.773	0.782	0.783	0.782	0.79	0.793	0.799	0.79	0.791	0.791	0.797	0.797
MDS - Uncertainty	0.813	0.796	0.797	0.791	0.791	0.802	0.81	0.81	0.812	0.806	0.81	0.811	0.808	0.81	0.809	0.81	0.807	0.803	0.802	0.806
Hard Hybrid - Certainty	0.748	0.763	0.762	0.77	0.78	0.782	0.778	0.783	0.787	0.792	0.788	0.78	0.782	0.787	0.787	0.787	0.79	0.792	0.791	0.787
Hard Hybrid - Uncertainty	0.804	0.797	0.787	0.773	0.736	0.744	0.745	0.74	0.745	0.752	0.744	0.743	0.751	0.758	0.761	0.762	0.76	0.762	0.76	0.767
BCS	0.734	0.756	0.765	0.771	0.772	0.77	0.762	0.762	0.762	0.761	0.763	0.767	0.77	0.761	0.755	0.748	0.75	0.749	0.756	0.756

LCP_PileEtFace_2012-01-12_192700, Head Track: 330, Speaker Track: 269, OCR Name: 52 HvH 228X228 SvS 218X218 SvH 373



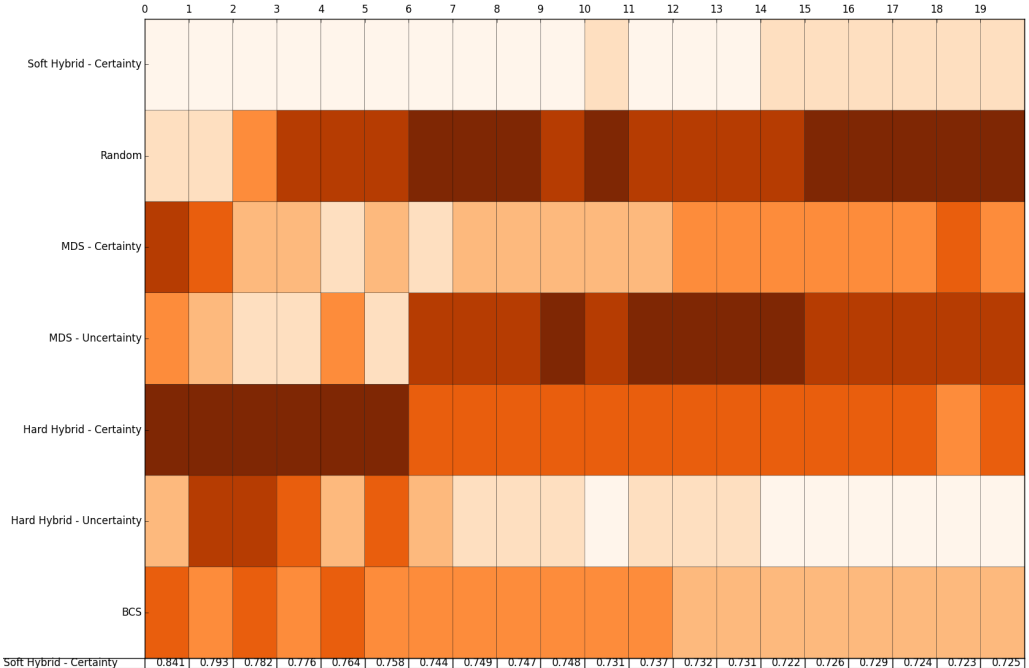
Soft Hybrid - Certainty	0.448	0.482	0.47	0.461	0.478	0.503	0.499	0.491	0.483	0.474	0.466	0.452	0.447	0.431	0.422	0.413	0.414	0.412	0.413	0.411
Random	0.531	0.443	0.426	0.448	0.459	0.466	0.441	0.446	0.444	0.445	0.437	0.432	0.416	0.419	0.42	0.419	0.421	0.426	0.424	0.42
MDS - Certainty	0.398	0.43	0.446	0.462	0.449	0.463	0.466	0.474	0.464	0.456	0.448	0.446	0.434	0.426	0.421	0.414	0.401	0.399	0.398	0.395
MDS - Uncertainty	0.494	0.453	0.485	0.472	0.481	0.489	0.474	0.469	0.469	0.453	0.436	0.431	0.426	0.416	0.412	0.397	0.388	0.389	0.376	0.374
Hard Hybrid - Certainty	0.559	0.497	0.486	0.497	0.514	0.509	0.497	0.511	0.509	0.507	0.511	0.499	0.494	0.479	0.482	0.475	0.474	0.467	0.464	0.455
Hard Hybrid - Uncertainty	0.497	0.515	0.492	0.464	0.458	0.451	0.462	0.467	0.471	0.453	0.452	0.44	0.442	0.436	0.426	0.423	0.425	0.426	0.421	0.426
BCS	0.639	0.578	0.534	0.526	0.501	0.49	0.49	0.478	0.477	0.474	0.476	0.476	0.475	0.474	0.466	0.47	0.461	0.458	0.455	0.446

LCP_PileEtFace_2012-01-19_192700, Head Track: 302, Speaker Track: 246, OCR Name: 43 HvH 222X222 SvS 191X191 SvH 290



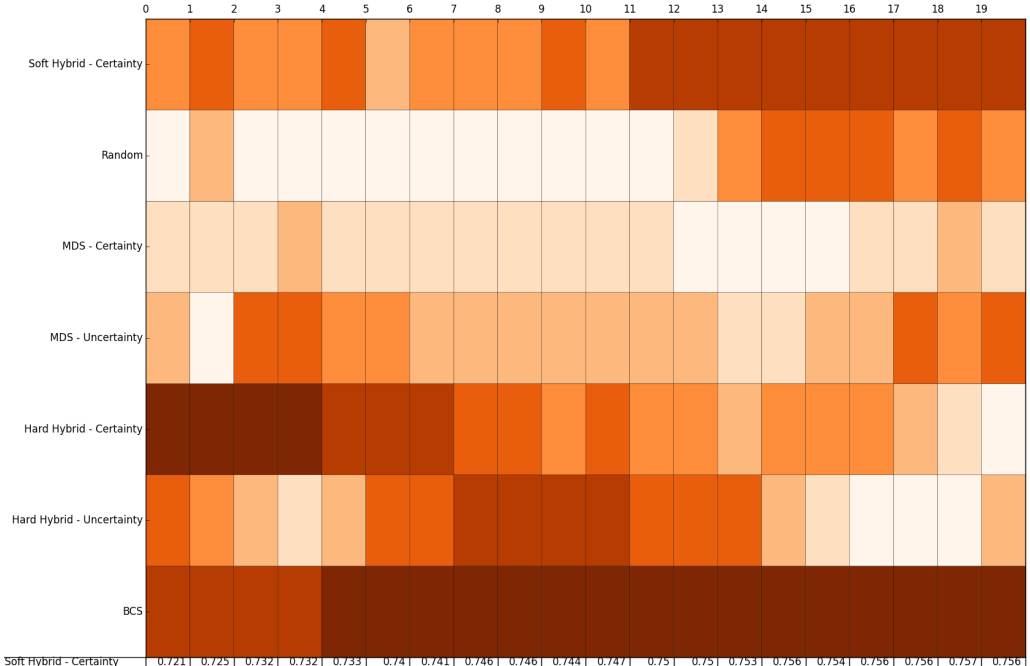
Soft Hybrid - Certainty	0.431	0.476	0.446	0.42	0.433	0.425	0.409	0.403	0.415	0.413	0.409	0.408	0.414	0.403	0.404	0.403	0.398	0.402	0.406	0.406
Random	0.45	0.382	0.324	0.308	0.279	0.283	0.287	0.288	0.29	0.29	0.293	0.301	0.312	0.322	0.326	0.325	0.331	0.336	0.333	0.333
MDS - Certainty	0.466	0.411	0.428	0.401	0.403	0.404	0.414	0.402	0.403	0.4	0.406	0.416	0.417	0.418	0.426	0.417	0.409	0.407	0.401	0.401
MDS - Uncertainty	0.507	0.51	0.438	0.423	0.412	0.395	0.383	0.378	0.373	0.372	0.373	0.373	0.384	0.375	0.377	0.378	0.376	0.376	0.38	0.381
Hard Hybrid - Certainty	0.369	0.338	0.349	0.325	0.324	0.328	0.347	0.339	0.345	0.348	0.361	0.361	0.364	0.356	0.361	0.364	0.365	0.369	0.371	0.378
Hard Hybrid - Uncertainty	0.361	0.401	0.375	0.374	0.339	0.346	0.373	0.38	0.376	0.359	0.371	0.368	0.364	0.365	0.363	0.362	0.358	0.348	0.352	0.355
BCS	0.41	0.408	0.401	0.387	0.362	0.363	0.368	0.361	0.354	0.345	0.35	0.356	0.365	0.372	0.371	0.373	0.375	0.376	0.378	0.379

LCP_PileEtFace_2012-01-26_192700, Head Track: 380, Speaker Track: 273, OCR Name: 50 HvH 223X223 SvS 207X207 SvH 349



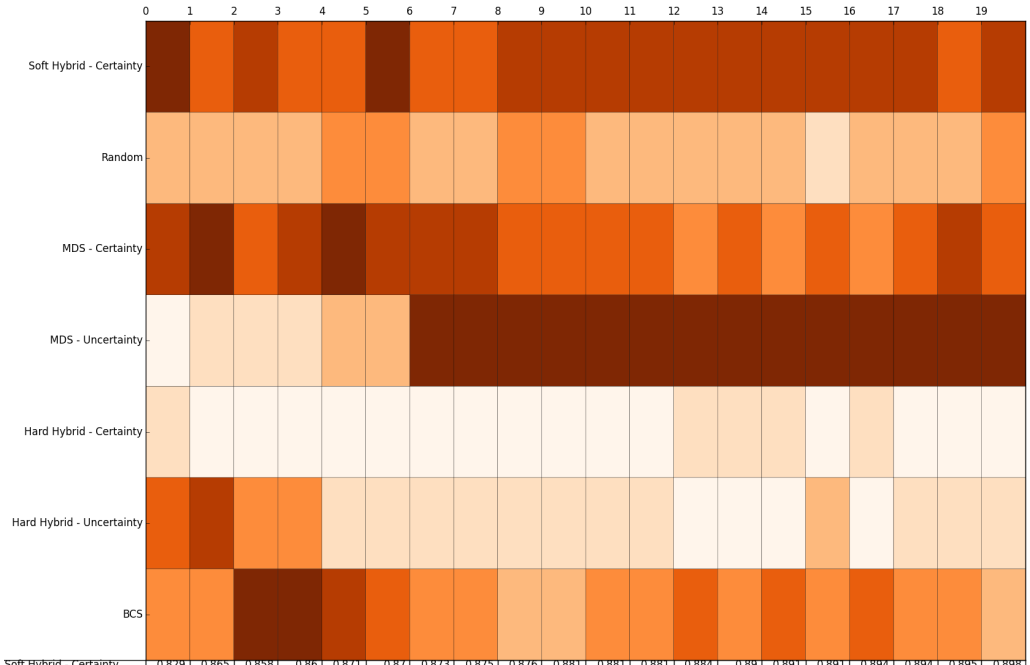
Soft Hybrid - Certainty	0.841	0.793	0.782	0.776	0.764	0.758	0.744	0.749	0.747	0.748	0.731	0.737	0.732	0.731	0.722	0.726	0.729	0.724	0.723	0.725
Random	0.834	0.765	0.731	0.71	0.715	0.71	0.691	0.693	0.699	0.695	0.691	0.693	0.699	0.684	0.689	0.686	0.684	0.687	0.682	0.683
MDS - Certainty	0.668	0.72	0.74	0.745	0.751	0.742	0.741	0.737	0.738	0.736	0.731	0.724	0.721	0.721	0.714	0.714	0.711	0.701	0.697	0.705
MDS - Uncertainty	0.716	0.745	0.755	0.754	0.741	0.745	0.711	0.71	0.705	0.684	0.691	0.679	0.679	0.676	0.681	0.686	0.693	0.69	0.693	0.695
Hard Hybrid - Certainty	0.635	0.689	0.686	0.706	0.706	0.708	0.719	0.716	0.722	0.719	0.715	0.713	0.705	0.704	0.706	0.698	0.7	0.698	0.7	0.697
Hard Hybrid - Uncertainty	0.729	0.716	0.72	0.724	0.743	0.735	0.732	0.739	0.745	0.744	0.742	0.735	0.73	0.73	0.731	0.734	0.736	0.735	0.736	0.727
BCS	0.713	0.73	0.73	0.729	0.727	0.736	0.732	0.729	0.738	0.735	0.727	0.723	0.728	0.726	0.721	0.721	0.723	0.724	0.72	0.722

LCP_TopQuestions_2012-01-25_232700, Head Track: 364, Speaker Track: 134, OCR Name: 11 HvH 204X204 SvS 102X102 SvH 310

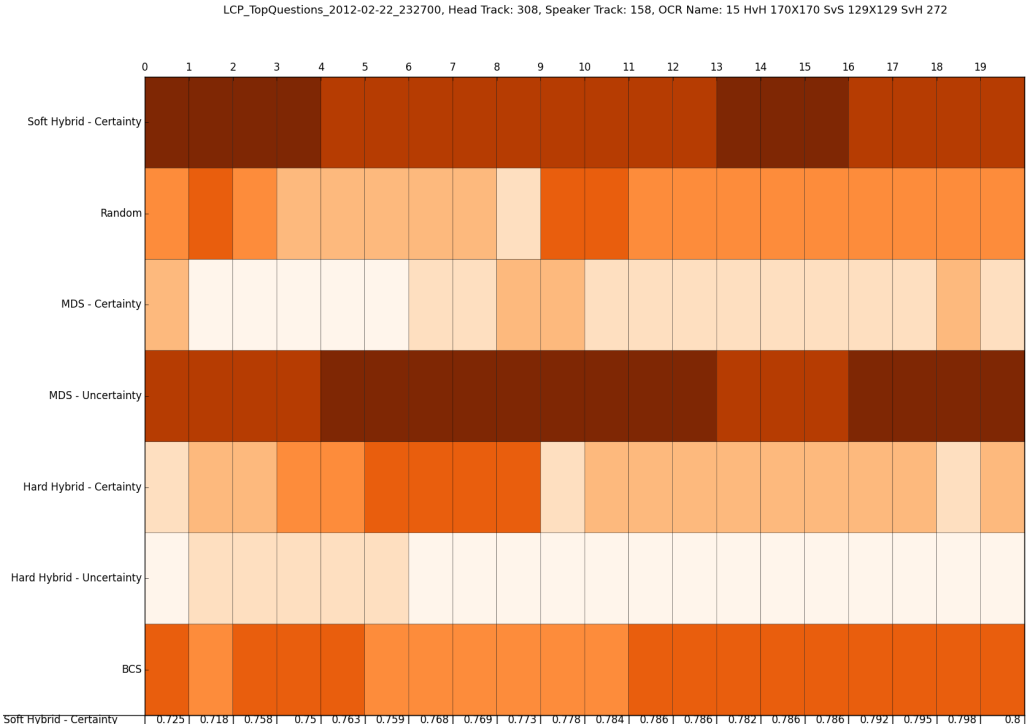


Soft Hybrid - Certainty	0.721	0.725	0.732	0.732	0.733	0.74	0.741	0.746	0.746	0.744	0.747	0.75	0.75	0.753	0.756	0.754	0.756	0.756	0.757	0.756
Random	0.764	0.74	0.753	0.753	0.753	0.758	0.772	0.767	0.767	0.771	0.763	0.759	0.757	0.755	0.757	0.758	0.759	0.761	0.762	0.763
MDS - Certainty	0.764	0.761	0.752	0.738	0.747	0.75	0.748	0.75	0.754	0.755	0.756	0.759	0.763	0.766	0.765	0.767	0.764	0.766	0.764	0.766
MDS - Uncertainty	0.757	0.762	0.73	0.724	0.733	0.736	0.745	0.748	0.749	0.751	0.753	0.755	0.754	0.757	0.761	0.759	0.761	0.76	0.763	0.762
Hard Hybrid - Certainty	0.697	0.712	0.694	0.709	0.724	0.732	0.738	0.741	0.746	0.746	0.747	0.751	0.752	0.757	0.757	0.758	0.76	0.763	0.764	0.767
Hard Hybrid - Uncertainty	0.719	0.738	0.733	0.741	0.738	0.736	0.739	0.737	0.741	0.744	0.745	0.751	0.752	0.754	0.758	0.762	0.765	0.768	0.766	0.763
BCS	0.706	0.718	0.724	0.717	0.719	0.723	0.724	0.723	0.724	0.722	0.724	0.728	0.732	0.734	0.731	0.732	0.735	0.735	0.736	0.736

LCP_TopQuestions_2012-02-14_232200, Head Track: 249, Speaker Track: 155, OCR Name: 13 HvH 133X133 SvS 127X127 SvH 272



Soft Hybrid - Certainty	0.829	0.865	0.858	0.86	0.871	0.87	0.873	0.875	0.876	0.881	0.881	0.881	0.884	0.89	0.891	0.891	0.894	0.894	0.895	0.898
Random	0.922	0.904	0.903	0.895	0.884	0.881	0.883	0.891	0.891	0.888	0.894	0.896	0.897	0.9	0.9	0.902	0.9	0.899	0.902	0.9
MDS - Certainty	0.856	0.834	0.867	0.859	0.866	0.872	0.871	0.874	0.877	0.883	0.888	0.895	0.897	0.893	0.895	0.895	0.895	0.896	0.895	0.899
MDS - Uncertainty	0.947	0.917	0.916	0.895	0.894	0.884	0.866	0.87	0.871	0.876	0.877	0.878	0.881	0.878	0.88	0.88	0.879	0.88	0.883	0.886
Hard Hybrid - Certainty	0.936	0.928	0.929	0.918	0.917	0.902	0.905	0.904	0.902	0.895	0.895	0.899	0.899	0.901	0.901	0.903	0.902	0.904	0.907	0.906
Hard Hybrid - Uncertainty	0.894	0.864	0.886	0.884	0.898	0.895	0.899	0.893	0.898	0.893	0.894	0.896	0.9	0.902	0.903	0.901	0.903	0.904	0.905	0.906
BCS	0.897	0.866	0.841	0.858	0.869	0.876	0.882	0.884	0.892	0.89	0.894	0.896	0.896	0.897	0.894	0.896	0.895	0.898	0.899	0.901



Soft Hybrid - Certainty	0.725	0.718	0.758	0.75	0.763	0.759	0.768	0.769	0.773	0.778	0.784	0.786	0.786	0.782	0.786	0.786	0.792	0.795	0.798	0.8
Random	0.759	0.749	0.785	0.808	0.802	0.801	0.805	0.805	0.807	0.801	0.803	0.803	0.806	0.805	0.81	0.811	0.809	0.809	0.813	0.813
MDS - Certainty	0.811	0.857	0.85	0.839	0.838	0.827	0.816	0.813	0.805	0.806	0.813	0.815	0.814	0.817	0.822	0.825	0.823	0.825	0.826	0.829
MDS - Uncertainty	0.752	0.736	0.763	0.758	0.76	0.755	0.761	0.759	0.764	0.766	0.773	0.779	0.783	0.789	0.791	0.789	0.79	0.793	0.793	0.796
Hard Hybrid - Certainty	0.819	0.81	0.785	0.787	0.793	0.778	0.782	0.791	0.798	0.806	0.805	0.807	0.808	0.809	0.815	0.818	0.819	0.822	0.827	0.828
Hard Hybrid - Uncertainty	0.824	0.831	0.829	0.82	0.822	0.825	0.82	0.823	0.823	0.822	0.825	0.824	0.826	0.827	0.83	0.831	0.829	0.828	0.83	0.831
BCS	0.759	0.785	0.782	0.779	0.789	0.786	0.795	0.805	0.805	0.805	0.805	0.8	0.802	0.799	0.797	0.8	0.801	0.802	0.803	0.806

CURRICULUM VITAE

Name Surname: Emre Demir

Place and Date of Birth: Zile, 02.08.1985

Adress: Tübitak Gebze Yerleşkesi, MAM Lojmanları, 31.Blok, D:5, Gebze/Kocaeli

E-Mail: ed.emre@gmail.com

B.Sc.: Computer Engineering, Ege University, 2009

Professional Experience and Rewards:

- 2014 April - Present, Turkcell Teknoloji, R&D department Product and Services division, as Developer.
  - Turkcell Self-Service(ongoing project), sending bulk SMS project, (Developed some parts of project in Java and leading the project).
  - Lawful interception module developed for Turkcell Multi-Media Message Service(MMS), (Java Web Services, Java Crypto Extension).
- 2011 May - 2014 April, The Scientific and Technological Research Council of Turkey, Institute of Cryptology and Electronics - TUBITAK UEKAE, Researcher & Software Engineer.
  - Worked on Electronic Key Management System, (Java Crypto Architecture, Java Crypto Extension, PKI, PKCS 11, Basic Cryptography).
  - Development of Java crypto library, national or NATO key generation algorithms, parity check algorithms etc. implemented (Java Crypto Architecture, Java Crypto Extension).
  - IFFMOD-5 key capability to Electronic Key Management System of Turkish Army.
- 2009 August - 2011 May, Defne Bilgi Islem, Istanbul Office at ITU Teknokent, Software Engineer in R&D Department.
  - Multi-Thread SMPP Server for NCell: High performance SMPP server developed. (Java, SMPP Stack, Sockets, Sonic MQ).
  - Queued Multi-Threaded Diameter Gateway: PayForMe upgrade for Avea (Java, Diameter Stack, Sockets, Linux).

- Multi-Thread Socket Server for IVR and DB Connection Pool, SMPP Client, Web Services, Back Ground Music Project, for Mobilink Pakistan (Java, Tomcat, SMPP Stack, Sockets, Linux).
- Web Services, SMPP Client, Malicious Call Barring Project for Azercell (Java, Tomcat, SMPP Stack, Linux).
- Making Videos From Web Pages, Research Project (FFMpeg API, C, Linux).
- Video Transcoding and Streaming, IP TV Research Project (FFMpeg API, C, Linux).

List of Publications and Patents:

PUBLICATIONS/PRESENTATIONS ON THE THESIS

▪ Unsupervised Active Learning for Video Annotation, International Conference of Machine Learning 2015, Active Learning Workshop, Lille - France.