



HAL
open science

Transferable belief model for hair mask segmentation

Cedric Rousset, Pierre-Yves Coulon, Michèle Rombaut

► **To cite this version:**

Cedric Rousset, Pierre-Yves Coulon, Michèle Rombaut. Transferable belief model for hair mask segmentation. ICIP 2010 - 17th IEEE International Conference on Image Processing, Sep 2010, Hong Kong, Hong Kong SAR China. pp.n.c. hal-00530156

HAL Id: hal-00530156

<https://hal.science/hal-00530156>

Submitted on 27 Oct 2010

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

TRANSFERABLE BELIEF MODEL FOR HAIR MASK SEGMENTATION

C. Rousset, P.Y. Coulon and M. Rombaut

GIPSA-Lab/Dis, Grenoble INP, UJF, CNRS
Domaine Universitaire – BP46 – 38402 St Martin d’Hères - France

ABSTRACT

In this paper, we present a study of transferable belief model for automatic hair segmentation process. Firstly, we recall the transferable Belief Model. Secondly, we defined for the parameters which characterize hair (Frequency and Color) a Basic Belief assignment which represents the belief that a pixel was or not a hair pixel. Then we introduce a discounting function based on the distance to the face to increase the reliability of our sensors. At the end of this process, we segment the hair with a matting process. We compare the process with the logical fusion. Results are evaluated using semi-manual segmentation references

Index Terms — Hair segmentation, Data Fusion, Transferable Belief Model

1. HAIR SEGMENTATION

Hair is an important feature of human appearance. Using hair in face recognition was considered until recently as a non stable parameter because hair aspect can be easily changed. However Hair may provide a cue for identification process. It can facilitate the process by providing parameters to describe a person and reducing possible matches.

Hair analysis has been very recently developed in scientific literature hence only few publications deal with this topic. Hair segmentation can result from head detection divide into a face and hair area [1] or from color model distribution [2][3]. The most relevant paper was published by Yacoob and Davis [4]. They used an automatic segmentation by region growing process based on a color model. Once the hair mask is segmented, they characterize the hair style and defined parameters (length, volume, color, etc...).

Here, we present our algorithm for the automatic detection and segmentation of hair (Fig.1) [5]. The approach consists in the automatic definition of markers which represent a raw segmentation in 3 classes: H , \bar{H} and U_n which represent respectively the hair pixels, non hair pixels and the *undefined* pixels. These markers are obtained from a frequential and a color analysis of the picture. The *undefined* pixels will be estimated during the matting process [6][7]. Then the hair mask is segmented according to the estimated alpha matte α .

The purpose of this study was to improve the fusion process in order to be more robust for positioning the markers for matting process. *Uncertainty* level is introduced taking into account ‘high level’ information like the distance to the face.

This paper is organized as follows. Section 2 presents the aim of our fusion data process. In section 3, we introduce transferable belief model for hair segmentation. Section 4 shows our results and we conclude in section 5.

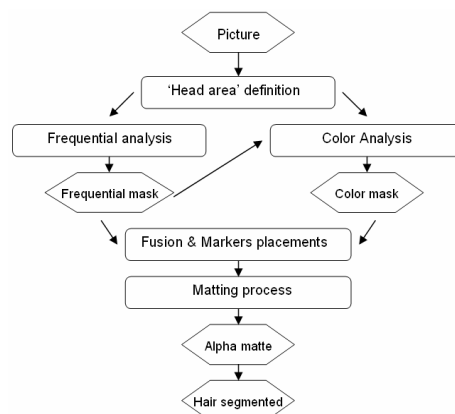


Fig 1. Representation of the sequence of the hair segmentation process

2. DATA FUSION

Data fusion combines data from multiple sources (sensors) and merge information in order to obtain better information (more accurate and more robust).

For hair segmentation, two features were defined (Frequency texture & Color) which can characterize hair position in a picture. We need to keep the common information obtained from both sensors in order to define the most probable hair position.

2.1. Basic Hair Data Fusion: Logical Fusion

Here we present the description of the trivial data fusion processed in the hair segmentation algorithm [5].

- Binary frequential map (F) was obtained after thresholding filtered luminance map with an isotropic Gaussian bandwidth filter.
- Binary color map (C) was obtained after thresholding the Cb & Cr Gaussian distribution from a color model. The color model is calculated from a window sample of the picture placed at the top of the face found by Viola & Jones process [8]. As we assume that hair information can be only present in the frequential map, the color model is calculated from the pixels inside the frequential mask.

Once binary frequential (F) and color (C) masks were computed, information taken from both masks was intersected following the Tab 1 rules:

For performing the matting treatment, H markers ($\alpha = 1$) are placed inside the fusion mask and \bar{H} markers ($\alpha = 0$) outside the frequential mask. The pixel which are in the frequential mask and

not in the color mask are not enough relevant to be characterized, they are considered as *Un undefined* and no markers were defined. These pixels are estimated during matting process. A better definition and localisation of these pixels should improve the matting process

Tab 1: Tab rules of Logical fusion

F	C		α
0	0	\bar{H}	0
0	1	\bar{H}	0
1	0	Un	\emptyset
1	1	H	1

3. TBM FOR HAIR SEGMENTATION

The Transferable Belief Model (TBM) developed by Smets [9] provides a model for the representation of quantified belief. We'll only recall some basics of the belief function theory. More details can be found in [10]

With the use of TBM, *Uncertainty Un* is introduced which represents a new possible state for a pixel that act as a substitute for the *Un undefined* pixels. So the definition of the pixels estimated during the matting treatment should be improved and we can guide the treatment with 'high level' information.

3.1. Belief Modelisation

Basic Belief Assignments (BBA) are defined based on each sensor to represent the knowing of the system. Let $\Omega = \{H \text{ (hair pixel)}, \bar{H} \text{ (non hair pixel)}\}$ denotes the finite set of possible events (values). The solution of the considered problem is unique and it is obligatory one of the element of Ω :

$$A_i \in 2^\Omega = \{\{\emptyset\}, \{H\}, \{\bar{H}\}, \{Un\}\} \quad (3.1)$$

A belief mass m is defined to each element of 2^Ω following the equation:

$$m: 2^\Omega \rightarrow [0,1], A \mapsto m(A) \quad (3.2)$$

For each sensor (Frequential and Color map), a BBA is associated to each events of Ω :

- The belief that a pixel is a hair pixel : $m(\{H\})$
- The belief that a pixel is a non hair pixel : $m(\{\bar{H}\})$
- The belief that the state of a pixel is uncertain : $m(Un)$
- The belief that the state of a pixel is unknown: $m(\emptyset)$. A BBA is normalized setting $m(\emptyset) = 0$

And for each sensor the BBA distribution verified the equation:

$$\sum_{A_i \in 2^\Omega} m(A_i) = 1 \quad (3.3)$$

Then we see that the TBM allows to model doubt in decision process. That is why it is well adapted to design a fusion approach to provide more reliable decision from different sensors of information.

Fig.2 represents the BBA model for $\Omega = \{H, \bar{H}\}$. The threshold values are determined for each descriptor.

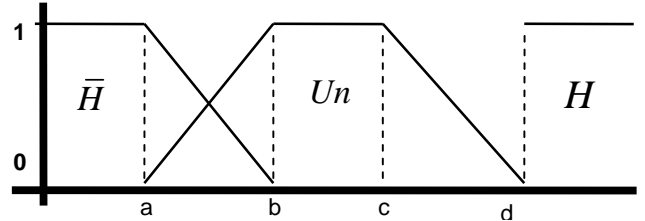


Fig 2. Representation of a BBA model for $\Omega = \{H, \bar{H}\}$.

3.2. Frequential BBA: m_F

From the obtained binary map, we have to create a descriptor that fits for our BBA definition. A diffusion process was performed based on an iterative mean filter on the frequential map. Pixels with the maximum value correspond to the pixels with the most reliability of being hair pixel.

The threshold values of the frequential BBA m_F are defined based on the mean of frequential binary mask repartition into an 8-connected neighborhood.

$$a=0 \quad b=1/9 \quad c=2/9 \quad d=3/9 \quad (3.4)$$

The values of the threshold are low due to the diffusion algorithm which leads to a strong decrease of the value of the frequential descriptor from the binary map. The minimum threshold a is fixed to zero because we assume that only the area with no frequency are totally defined as non hair pixel. The value of b (respectively c & d) corresponds to the mean ratio to 1 (respectively 2 & 3) pixels equal to one in the neighborhood.

3.3. Color BBA: m_C

As we can't create the color descriptor from the binary color map we use a model based on the Gaussian distribution of the color model (cf 2.1). We consider the color Gaussian distribution as the density of probability for a pixel to be a hair pixel.

The threshold values of the color BBA m_C are then defined based on the standard deviation of Gaussian distribution.

$$a=4\sigma \quad b=3\sigma \quad c=2\sigma \quad d=\sigma \quad (3.5)$$

Thus the pixels near (respectively far) the mean of the Gaussian distribution are considered as H (respectively \bar{H}) pixels. The others are defined as Un pixels.

3.4. Combination

From the two given BBAs m_F and m_C defined for the same Ω , combination rules are used in order to fusion the information from the different sources. The conjunctive rule of combination (CRC) is used to achieve a new BBA defined on 2^Ω (Tab.2). The obtained BBA m_{FC} represents the global belief distribution which will facilitate the decision of the state. After combination, the empty set can have a belief non null. The belief on $m_{FC}(\emptyset)$ is transferred to $m_{FC}(Un)$ following the Yager rule [10].

Tab 2: m_{FC} tab of CRC of m_F and m_C

$m_C \backslash m_F$	H	\bar{H}	Un
H	H	\emptyset	H
\bar{H}	\emptyset	\bar{H}	\bar{H}
Un	H	\bar{H}	Un

3.5. Discounting BBA

The two descriptors are based on hair parameter without taking into consideration ‘high level’ information which can improve the reliability of each event of Ω . One particularity of hair is their position, even if their position is highly variable and it can’t be exactly define where they are, we can assume that hair is around the face. A discounting BBA is defined based on a distance function obtained from a localization of the face mask.

3.5.1. Distance Face map

As for the hair color model, a skin model color was calculated from a sample window of a part of the face. The window is obtained from the use of eyes detection algorithm combined with the face detection of Viola & Jones. A sample window around the nose is generally selected. It leads to a better skin sample as we have some luminance aspect similar to the entire face on a small sample.

Once the model is calculated, the picture is threshold to obtain a binary map of pixel similar to skin. In order to avoid the neck, shoulder or more skin pixels and background pixels similar to face color, a labeling approach is performed to select the biggest area in the Viola & Jones window.

Once the Face mask obtained, the distance map is calculated taking into consideration that hair is generally outside the face but for example some locks can be in front of the face.

So two distance maps D_{in} and D_{out} are created:

- D_{in} is based on the distance outside the face
- D_{out} is based on the distance inside the face.

3.5.2. Discounting BBA

In TBM, the reliability reflects the user opinion about the ‘value’ of the sensor reading. The idea is to weight most heavily the opinion of the best sensor and conversely for the less one.

The coefficient $D = D_{in} + D_{out}$ (Fig.3) can be defined as the degree of relevance of the sensors based on the distance to the face. The more the sensor is reliable (Low value of D), the less the BBA will be affected.

We defined m_D the discounting BBA based on the *Distance-based* model of Denoeux [11]:

$$\begin{aligned} m_D(H) &= (1-D).m_{FC}(H) \\ m_D(\bar{H}) &= D.m_{FC}(\bar{H}) \\ m_D(Un) &= m_{FC}(Un) + D.m_{FC}(H) + (1-D).m_{FC}(\bar{H}) \end{aligned} \quad (3.6)$$

With the discounting BBA, the zones of certainty of H and \bar{H} are decreased and the zone of *uncertainty* Un is increased according to the distance map

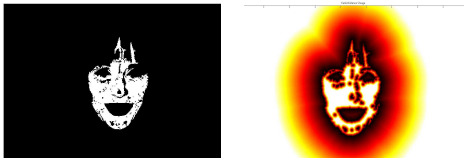


Fig 3. Representation of the distance maps of the face: The face mask segmented (left) and the distance map obtained D (right). The distance is normalized between 0 (black) to 1 (white).

3.6. Decision

The finality of a fusion data process is to take a decision about the possible state of the system. For TBM, the decision is generally taken for the events with the higher pignistic probability [8][9].

As we consider the uncertainty as a possible belief, the decision is simplified. The state of each pixel is decided from the belief function with the max value between $m(H)$, $m(\bar{H})$ and $m(Un)$. So three binary masks are created: one for hair pixels, one for non hair pixels and the last one for *uncertain* pixels. These masks are cleaned performing morphological opening in order to remove small artifacts.

4. EXPERIMENTAL RESULTS

We use a part of the CALTECH base [12]. This is a set of 450 images of 27 persons taken under several lighting conditions and backgrounds. We use only women models because women have more different hairstyles which include a part of men hair style and we avoid facial hair problems (beard, goatee, and mustache).

4.1. Methods Comparison

To validate the use of TBM for improve the hair mask segmentation, we compare the hair mask segmented after matting process initialized by the different approaches of automatic markers placement: Logical fusion (LF), TBM (TBM) and TBM with discounting BBA (TMBD).

Moreover in order to test the ‘high level’ information, defined by the distance to the face, on the Logical approach, the markers obtained from the Logical fusion are weighted (LFD) based on the distance map (cf 3.5.1). The values of α (Tab.1) are balanced from 1 (respectively 0) to [0.5,1] (respectively [0,0.5]) .

4.2. Result evaluation

A ground truth database is created using hair masks obtained after matting process initialized with markers placed manually. We measure the performance of our algorithm based on 2 criteria: false rejection rate (FRR) and false acceptance rate (FAR):

- FRR corresponds to a H pixel segmented as a \bar{H} pixel
- FAR corresponds to a \bar{H} pixel segmented as a H pixel

4.3. Performance Comparison

Results (Tab.3) show that the LFD approach leads to bad results. Weighting the markers is not adapted to matting process. So matting process need a good localisation of the initial markers rather than weighted values.

Results show also that LF and TBM approaches have similar results. Even if the two methods are different, the markers position obtained leads to same alpha matte estimation. The LF results are slightly different from previous paper because we compare only the raw result of the logical fusion without improvement.

Finally, the better results are obtained with TMBD. The hair segmentation is improved with a process more accurate and especially more robust with a standard deviation of false rejection down to 4%. These results show that for all pictures the entire hair mask is well segmented. Nevertheless the rate of false acceptance is always high and close to rate obtained by LF and TBM approaches but the qualitative result is quite different (Fig.4).

Tab 3: Results obtained for 40 pictures of women models

	FRR	σ (FRR)	FAR	σ (FAR)
LF	6,7 %	16,3 %	40,62 %	21,63 %
LFD	22,73 %	20,32 %	45,38 %	9,59 %
TBM	7,32 %	15,4 %	40 %	26,67 %
TBMD	2,68 %	4,34 %	37,63 %	20,96 %



Fig 4: Comparison between TBM without (top) and with (bottom) the discounting BBA. We obtain a similar FAR around 170% but we achieve a better localization of the over segmented pixels with the discounting BBA.

4.3. Discounting Improvement

In TBMD approach, as the uncertain zone has been increased for middle range distance of the map face, the over segmented pixels are in a peripheral area around the face. So we should enhance our discounting approach to limit the over segmentation (*Tab.4*)

- In TBMD2, the discounting function was applied only on the H hair pixel. We didn't apply the discounting function to \bar{H} pixels in order to keep strong these markers especially for the pixels near the face and neck. We achieve to decrease the false acceptance rate down to only 12% but we increase the false rejection rate to 13%.
- In TBMD3, the face map was redefined and takes into account a part of the neck and the discounting function was applied on both H & \bar{H} . We achieve similar results, the over segmented pixels are just moved away of the new face map.

So we can control the over segmented pixels. As for further development, an algorithm of decision will be developed in order to guide the segmentation. In function of the picture we should segmented the best hair mask or reduced the over segmented pixels.

5. CONCLUSION

A data fusion process is developed for hair mask segmentation. The use of Transferable Belief Model improves the fusion of the Frequential and Color sensors and leads to more probable position of hair pixels. TBM is combined with a discounting BBA based on the distance to the face in order to take into account 'high level information' of hair position. Moreover the addition of the *Uncertain* pixels which represent neither hair or non hair pixels improves the matting process.

The experiment shows that the proposed method is effective and gives better results. We achieve to segment generally almost all the hair mask with a very low rate of false rejection. And even if the rate of false acceptance is still high, we achieve a better definition and localization of the over segmented pixels.

Tab 4: Results obtained for 40 pictures of women models

	FRR	σ (FRR)	FAR	σ (FAR)
TBMD	2,68 %	4,34 %	37,63 %	20,96 %
TBMD2	13,6 %	13 %	12,9 %	9,1 %
TBMD3	4,6 %	5,4 %	30,7 %	15,4 %

6. REFERENCES

- [1] M.Kampmann, "Segmentation of a Head into Faces, Ears, Neck and Hair for Knowledge-based analysis synthesis coding of videophone sequence", IEEE ICIP, 1998
- [2] H.Wu, Q.Chen and M.Yachida, "Face detection from color images using a fuzzy pattern matching method", IEEE Trans. PAMI, 557-563, 1999.
- [3] Z. Zhang, H.Gunes and M.Piccardi, "An Accurate Algorithm For Head Detection Based On XYZ and HSV Hair and Skin Color Model", IEEE ICIP, 2008
- [4] Y.Yacob and L.S.Davis, "Detection and analysis of hair", IEEE Trans. Pattern Analysis and Machine Intelligence, 1164-1169, 2006.
- [5] C.Rousset and P.Y.Coulon, "Frequential and Color Analysis for Hair mask Segmentation", IEEE ICIP, 2276-2279, 2008
- [6] D.Lischinski et Y.Weiss A.Levin, "A closed form solution to natural image matting", IEEE Computer Vision and Pattern Recognition, 2006.
- [7] D.Lischinski A.Levin and Y.Weiss, "Compositing digital image", SIGGRAPH, 689-694, 2004.
- [8] P.Viola and M.Jones, "Compositing Rapid object detection using a boosted cascade of simple features", IEEE Computer vision and Pattern recognition, 511-518, 2001.
- [9] Ph Smets and R. Kennes, "The Transferable Belief Model", Artificial Intelligence, 66(2), 191-234, 1994.
- [10] E. Ramasso, "Reconnaissance de Séquence d'Etat par le Modèle des Croyances Transférables", PhD Thesis, 2007.
- [11] T. Denoeux, "A k-Nearest Neighbor Classification Rule Based on Dempster-Shafer Theory", IEEE Trans. On Systems Man and Cybernetics, 804-813, 1995.
- [12] <http://vision.caltech.edu/html-files/archives.html>



Fig 5: Examples of good hair segmentation. Original picture (left) and segmentation results in red (right)