
Face classification applications

s1700260

Abstract

The face is perhaps the part of human body that carries the majority of information about a person. By analyzing the different attributes of a face we can make inferences about that person, e.g., gender, age, mood and more. In this review, I present a set of approaches used in computer vision to classify human faces and their applications. Moreover, I provide an analysis of the features used to perform the classification tasks along with their advantages and disadvantages. Finally, I discuss about the variations of the approaches and further work.

1. Introduction

Human classification is a common task performed in a multitude of scenarios. Its objective varies depending on the domain where it is performed, but in many cases it is meant to make inferences about a person of a given class. The range of applications varies from entertainment to surveillance including health, sports and more. The classification can be performed as simply as using a single feature of a subject or using more complex systems that analyse the correlation between different sets of features. No matter the way a system classifies people, one of the key aspects of every system is the data it uses; face is a part of every person that provides a lot of information.

As mentioned, the human face is a source for vast information. Depending on the features to analyze and the way they are analyzed, a subject can be classified into one group or another. For this reason it is important to have a good understanding of the domain where the classification is going to be performed in order to get proper results. For instance, in a system that is meant to determine the mood of a person, it makes more sense to use the facial expressions rather than the hair or eye color.

For the common eye, it is easy to identify a set of features from almost any object (faces included). However, in computer vision this is a non-trivial task. Several general purpose feature detectors from images have been developed e.g. (Harris & Stephens, 1988) and (Lowe, 1999), as well as domain specific ones: (Chow & Li, 1993) and (Colmenarez et al., 1999). No matter their domain, the ultimate goal of these approaches is to provide information in a meaningful way for a system, so this is able to perform a given task with good results. In classification, some alternatives are able to extract features and categorize faces at once (Fleming &

Cottrell, 1990).

The information extracted from faces can be used to feed a variety of systems for different purposes. Many of the alternatives are based on machine learning methods, being neural networks a popular choice for tasks like face detection (Rowley et al., 1998), face recognition, (Lawrence et al., 1997) and face verification (Sun et al., 2013); all of them intrinsically classification tasks.

2. Face information

Even though there are approaches that address the problem of feature detection and they have been proven to give good results, it is important to understand the information that we can get from the human face. This way it would be possible to get more insights about how some systems leverage the information from a face and how they can be improved or modified to perform other tasks. Here, I will focus in the geometry and color information provided by faces.

2.1. Geometry

Geometry-wise, there are two important characteristics in the human face. First, it is well structured, meaning that the positions of its different components have little variation from subject to subject; the mouth is located in the bottom, the eyes close to the top, and so on. Some applications take advantage of this property to synthesize data to train neural networks (Hu et al., 2016), copy a makeup style from one face to another (Guo & Sim, 2009), or to simply replace faces in photographs for entertainment purposes (Bitouk et al., 2008). Moreover, the well defined structure of a face along with its enormous differences to other elements make possible systems for face detection (Zhu et al., 2006)

In the other hand, the variability of the face components from subject to subject (shape, relative position, relative size, etc.) is high enough that it is possible to have approaches for face recognition (Ahonen et al., 2004) which can be used in applications for authentication or reidentification. In addition, the components of a face in a single subject are not static, (their shapes can vary) which is clearly evident in facial expressions, an attribute that can be used to recognize moods or emotions (Adolphs, 2002). This characteristic has been used in systems aimed to facilitate human-computer interaction (Yacoob & Davis, 1996).

Other than facial expressions, subjects can modify the geometry of their faces by using makeup, wearing accessories like glasses and earrings, or even by changing their hair style. This situation evidently increases the variability of

a single face, which supposes a challenge for systems that rely in the geometry information of faces. Some alternatives have addressed this problem by performing the correlation between the raw face and the modified one (Guo et al., 2014). Other approaches have focused their efforts in the detection of such accessories (Jing & Mariani, 2000) and (Wu et al., 2004)

2.2. Color

In computer vision, one of the salient characteristics of color is the ability to be represented in a multitude of formats. Thanks to this property, it is possible to convert the color information in a picture from the default color space (usually in the red, green and blue channels) to a more suitable format depending on the application domain. This is leveraged by a multitude of systems for different purposes including explicit image detection (Basilio et al., 2011) and face detection (Garcia & Tziritas, 1999)

The human skin color in digital images has been widely studied. One of the major developments thrown by previous researches in this area is the determination of the proper color space to study the human skin. Even though that a proper analysis of a color space can give good results for applications over human skin color (Albiol et al., 2001), YCbCr color space has demonstrated to give the best results. The main reason is that for skin detection, the majority of systems discard the luminance of the image (Y component) as stated in (Vezhnevets et al., 2003)

In the human face not only the skin provides color information, other components like the eyes and hair provide information in regards to that matter. The analysis of the color of these elements has proven to be useful in applications like gaze tracking (Kim & Ramakrishna, 1999) and person recognition and image indexing (Yacoob & Davis, 2006).

Like geometry, face color can be prone to changes too. Here the changes can also depend on the behaviour of the subjects e.g. the use of makeup. However, most of the color variability in an image comes from the configuration of the device to capture the scene (exposure time, shutter speed, etc.), and the environment setup (lighting, shadows, etc.). Nonetheless, skin color variability not only brings problems, it can be leveraged to make health related analysis (Tsumura et al., 2003)

3. Classification applications with faces

There is a wide range of applications that lay their foundations in the classification task using images with faces. Based on the geometry and color information mentioned earlier, I will describe and analysis three systems that uses that kind of data. The fist system uses geometry information represented as histogram of gradients (HOG) for face recognition (Déniz et al., 2011). The second system detect faces based on skin region detection using color tranformations (Hsu et al., 2002). Finally, the third systems uses both

geometry and color information to determine the presence of makeup in faces (Chen et al., 2013).

3.1. Face recognition using HOG

As stated by (Déniz et al., 2011) HOG is a technique founded over the number of edge orientations occurrences within the local vicinity in an image. In their work, the authors propose a method based on the previous project by (Albiol et al., 2008). The main differences with the previous work lies in the way HOG is performed, a combination of classifiers, and the databases used to ran the experiments. The different way to perform HOG is aimed to reduce the possible dependency on the facial landmarks localization of previous results. A facial landmark is a salient feature within a face, in other words, an interest point.

There are two proposed ways to perform HOG in a face image. The first one lies in the location of the facial landmarks as well, however, the key point here is redundancy. Once the facial landmarks are localized, HOG is performed over each one of them at different patch sizes. That means that the bigger the patch size, the higher the redundancy of the retrieved information by HOG as seen in Figure 2 in (Déniz et al., 2011). In the second method, an image face is simply divided in non overlapped patches over which HOG is executed.

As mentioned by the authors, the redundancy of the first method may cause overfitting if the classifier does not have a mechanism for feature selection. On the other hand, the second alternative could prevent that situation due to its non overlapping nature. However, some amount of redundancy may be necessary to avoid the other undesired effect in machine learning models: underfitting. In fact, some HOG implementations can be configured to create a certain amount of redundancy by adding normalization over overlaped blocks of cells.

Another option that could help to increase the robustness of the feature descriptor is the combination of several types of features. Even though, the authors experimented combinations of HOG with Principal Component Analysis (PCA) and Latent Dirichlet Allocation (LDA), other methods like Local Binary Patterns may give good results. In fact, an alternative that combines HOG and LBP has proven to perform well in adverse situations like occlusion as seen in (Wang et al., 2009).

The proposed method to make the recognition of a face is based in the idea of weak classifiers, something that has proven to be good in several scenarios including the very same face recognition task (Yang et al., 2004). Other machine learning method that can perform well in this kind of problems is neural networks as demonstrated in (Parkhi et al., 2015). However, the success of a neural network is highly tied to a big amount of data for training process.

Finally, the datasets used to perform the experiments presented considerable differences from one to another. As seen in Figure 4 of their work, the variance of the facial

landmarks is quite different from one database to the other. This is not a minor issue since the performance of the method varies according to the dataset used. There is no much that can be done here since each dataset has their own characteristics, some of them were even built for a specific purpose. However, the naive alternative is to combine all the databases to get a more robust classifier.

Even though the proposed method does not erupt in the computer vision field, it was able to improve the results of its antecesor. The main point in this solution is the way of extract face features; it uses the same mechanism of the previous approach, but its execution is different. As seen in many scenarios, a proper definition of features can lead to expected results and verify the validity of an hypothesis, something that can was accomplished in this work. Futher work could be aimed to verify the performance of the feature extraction methodology in other tasks like face detection.

3.2. Face detection in color images

The approach for face detection proposed by (Hsu et al., 2002) is highly focused in the analysis of color in images. The solution is motivated by the lack of previous color-based alternatives that are robust enough to detect skin colors in different light conditions and complex background. The proposed method relies in two main components. The first one is aimed to detect face candidates based on some color transformations performed in an image. The second element is focused in the validation of the face candidates by detecting facial features.

The authors are aware of the challenges related to the definition of skin tones in color images due to the lighting conditions, and the previous work in the field. Based on this knowledge, the solution starts by compensating the lighting condition by using a reference white area, which allows to reduce the number of false positives. As demonstrated in other related works, YCbCr is the preferred color space when it comes to human skin detection. However, unlike other solutions where the skin pixels were detected by simply checking their values against a predefined range, the algorithm performs a non-linear transformation the helps to reduce the false positives as well.

Once the face candidate region is detected, it is analyzed to find facial features. This analysis relies in the distinctive chroma an luminance components of eyes and mouth. In summary, the solution creates an intensity map that highlights the regions where eyes and mouth are located. Such intensity map is created through a sequence of non linear transformation of the image components in the YCbCr color space. Once the three regions are detected, all the triangle zones are computed and filtered under a threshold to finally detect a face.

As described, this solution makes an intense analysis of the color in an image. However, due to the lack of semantic analysis, some anthropomorphic areas can be detected as human faces e.g. toys. A further work could consider an

analysis of the context in the image to reduce the false positives and improve results. Finally, since the approach analyzes the luminance in an image, it could be used to detect fake or mounting images.

3.3. Automatic Facial Makeup Detection with Application in Face Recognition

Face recognition is a well studied task, there are many geometry-based approaches that have been proved to provide good results under certain conditions. However, in the real world there are many variables that can reduce the performance of such solutions. Makeup is one of those variables that challenges face recognition. The work proposed by (Chen et al., 2013) is founded over the lack of previous solutions that address the problem of makeup and given that using it is a common habit among a great percentage of the population.

In their work, they propose a method that combines a set of color, geometric and texture features aimed to detect the presence of makeup in a face. Moreover, the solution is extended to be used in face recognition. As stated in their work, the use of makeup changes the perceptual appearance of people by highlighting characteristics and reducing flaws. This has been proved to have a huge impact in the performance of some recognition systems. For this reason, it is neccesary to redefine the set of features used to feed those solutions.

The main contribution of this work is the redefinition of the face features in order to consider the changes in shape, color and texture due to the use of makeup. The main objective is to provide a set of descriptors that clearly determine the presence or absence of makeup, so a face recognition system can be tuned properly to maintain its performance under such conditions.

Unlike other works that make use of the YCbCr color space when it comes to human skin, here they use the HSV color space (Hue, Saturation, Value) that highlights makeup regions in the saturation channel. The solution proposes a definition of color descriptors by combining the local and global features. This is a good idea since it adds some redundancy but keeps fidelity over all images. Shape and texture descriptors are simpler but they still contribute to increase the robustness of the feature vector.

Their experiments proved to give good results, however since the work was highly focused in mid-age female image samples, further work could address a more general solution including other segments of the population. Moreover, since the solution used almost perfectly front face images, further work can also include profile pictures.

4. Conclusions and further work

As seen, the human face is a source of vast information and there are several classification applications that rely in that data. One of the key aspects of these applications is the way they define a face in terms of its features. Depending in the

domain of the application some features are more useful than others. But the solutions go beyond that; it is not about the kind of information used to address a problem, the way it is processed and presented is very relevant in the success of an approach.

In addition, it is clearly evident that a given task can be addressed in different ways. Nonetheless, another important component of every approach is its foundation in previous work, this way new solutions can go beyond what have been proposed previously. In the case of applications with human faces, there is a lot of work done in regards to analyse their information and make use of them. However, the development does not seem to stop and every time new alternatives are proposed. Further reviews could be aimed to investigate what kind of information from human faces could be retrieved from sources other than 2D images, including Laser Imaging Detection and Ranging (LIDAR) and medical imagery, and the applications that can be done out of that.

References

- Adolphs, Ralph. Recognizing emotion from facial expressions: psychological and neurological mechanisms. *Behavioral and cognitive neuroscience reviews*, 1(1):21–62, 2002.
- Ahonen, Timo, Hadid, Abdenour, and Pietikäinen, Matti. Face recognition with local binary patterns. *Computer vision-eccv 2004*, pp. 469–481, 2004.
- Albiol, Alberto, Torres, Luis, and Delp, Edward J. Optimum color spaces for skin detection. In *Image Processing, 2001. Proceedings. 2001 International Conference on*, volume 1, pp. 122–124. IEEE, 2001.
- Albiol, Alberto, Monzo, David, Martin, Antoine, Sastre, Jorge, and Albiol, Antonio. Face recognition using hog-ebgm. *Pattern Recognition Letters*, 29(10):1537–1543, 2008.
- Basilio, Jorge Alberto Marcial, Torres, Gualberto Aguilar, Pérez, Gabriel Sánchez, Medina, L Karina Toscano, and Meana, Hector M Perez. Explicit image detection using ycbcr space color model as skin detection. *Applications of Mathematics and Computer Engineering*, pp. 123–128, 2011.
- Bitouk, Dmitri, Kumar, Neeraj, Dhillon, Samreen, Bel-humeur, Peter, and Nayar, Shree K. Face swapping: automatically replacing faces in photographs. *ACM Transactions on Graphics (TOG)*, 27(3):39, 2008.
- Chen, Cunjian, Dantcheva, Antitsa, and Ross, Arun. Automatic facial makeup detection with application in face recognition. In *Biometrics (ICB), 2013 International Conference on*, pp. 1–8. IEEE, 2013.
- Chow, Gloria and Li, Xiaobo. Towards a system for automatic facial feature detection. *Pattern Recognition*, 26(12):1739–1755, 1993.
- Colmenarez, Antonio, Frey, Brendan, and Huang, Thomas S. Detection and tracking of faces and facial features. In *Image Processing, 1999. ICIP 99. Proceedings. 1999 International Conference on*, volume 1, pp. 657–661. IEEE, 1999.
- Déniz, Oscar, Bueno, Gloria, Salido, Jesús, and De la Torre, Fernando. Face recognition using histograms of oriented gradients. *Pattern Recognition Letters*, 32(12):1598–1603, 2011.
- Fleming, Michael K and Cottrell, Garrison W. Categorization of faces using unsupervised feature extraction. In *Neural Networks, 1990., 1990 IJCNN International Joint Conference on*, pp. 65–70. IEEE, 1990.
- Garcia, Christophe and Tziritas, George. Face detection using quantized skin color regions merging and wavelet packet analysis. *IEEE Transactions on multimedia*, 1(3):264–277, 1999.
- Guo, Dong and Sim, Terence. Digital face makeup by example. In *Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on*, pp. 73–79. IEEE, 2009.
- Guo, Guodong, Wen, Lingyun, and Yan, Shuicheng. Face authentication with makeup changes. *IEEE Transactions on Circuits and Systems for Video Technology*, 24(5):814–825, 2014.
- Harris, Chris and Stephens, Mike. A combined corner and edge detector. In *Alvey vision conference*, volume 15, pp. 10–5244. Manchester, UK, 1988.
- Hsu, Rein-Lien, Abdel-Mottaleb, Mohamed, and Jain, Anil K. Face detection in color images. *IEEE transactions on pattern analysis and machine intelligence*, 24(5):696–706, 2002.
- Hu, Guosheng, Peng, Xiaojiang, Yang, Yongxin, Hospedales, Timothy, and Verbeek, Jakob. Frankenstein: Learning deep face representations using small data. *arXiv preprint arXiv:1603.06470*, 2016.
- Jing, Zhong and Mariani, Robert. Glasses detection and extraction by deformable contour. In *Pattern Recognition, 2000. Proceedings. 15th International Conference on*, volume 2, pp. 933–936. IEEE, 2000.
- Kim, Kyung-Nam and Ramakrishna, RS. Vision-based eye-gaze tracking for human computer interface. In *Systems, Man, and Cybernetics, 1999. IEEE SMC'99 Conference Proceedings. 1999 IEEE International Conference on*, volume 2, pp. 324–329. IEEE, 1999.
- Lawrence, Steve, Giles, C Lee, Tsoi, Ah Chung, and Back, Andrew D. Face recognition: A convolutional neural-network approach. *IEEE transactions on neural networks*, 8(1):98–113, 1997.

Lowe, David G. Object recognition from local scale-invariant features. In *Computer vision, 1999. The proceedings of the seventh IEEE international conference on*, volume 2, pp. 1150–1157. Ieee, 1999.

Parkhi, Omkar M, Vedaldi, Andrea, Zisserman, Andrew, et al. Deep face recognition. In *BMVC*, volume 1, pp. 6, 2015.

Rowley, Henry A, Baluja, Shumeet, and Kanade, Takeo. Neural network-based face detection. *IEEE Transactions on pattern analysis and machine intelligence*, 20(1):23–38, 1998.

Sun, Yi, Wang, Xiaogang, and Tang, Xiaoou. Hybrid deep learning for face verification. In *Proceedings of the IEEE International Conference on Computer Vision*, pp. 1489–1496, 2013.

Tsumura, Norimichi, Ojima, Nobutoshi, Sato, Kayoko, Shiraishi, Mitsuhiro, Shimizu, Hideto, Nabeshima, Hirohide, Akazaki, Syuuichi, Hori, Kimihiko, and Miyake, Yoichi. Image-based skin color and texture analysis/synthesis by extracting hemoglobin and melanin information in the skin. *ACM Transactions on Graphics (TOG)*, 22(3):770–779, 2003.

Vezhnevets, Vladimir, Sazonov, Vassili, and Andreeva, Alla. A survey on pixel-based skin color detection techniques. In *Proc. Graphicon*, volume 3, pp. 85–92. Moscow, Russia, 2003.

Wang, Xiaoyu, Han, Tony X, and Yan, Shuicheng. An hog-lbp human detector with partial occlusion handling. In *Computer Vision, 2009 IEEE 12th International Conference on*, pp. 32–39. IEEE, 2009.

Wu, Bo, Ai, Haizhou, and Liu, Ran. Glasses detection by boosting simple wavelet features. In *Pattern Recognition, 2004. ICPR 2004. Proceedings of the 17th International Conference on*, volume 1, pp. 292–295. IEEE, 2004.

Yacoob, Yaser and Davis, Larry S. Recognizing human facial expressions from long image sequences using optical flow. *IEEE Transactions on pattern analysis and machine intelligence*, 18(6):636–642, 1996.

Yacoob, Yaser and Davis, Larry S. Detection and analysis of hair. *IEEE transactions on pattern analysis and machine intelligence*, 28(7):1164–1169, 2006.

Yang, Peng, Shan, Shiguang, Gao, Wen, Li, Stan Z, and Zhang, Dong. Face recognition using ada-boosted gabor features. In *Automatic Face and Gesture Recognition, 2004. Proceedings. Sixth IEEE International Conference on*, pp. 356–361. IEEE, 2004.

Zhu, Qiang, Yeh, Mei-Chen, Cheng, Kwang-Ting, and Avidan, Shai. Fast human detection using a cascade of histograms of oriented gradients. In *Computer Vision and Pattern Recognition, 2006 IEEE Computer Society Conference on*, volume 2, pp. 1491–1498. IEEE, 2006.