

Joint Inference for Extracting Text Descriptor from Triage Images of Mass Disaster Victims

Niyati Chhaya
CSEE, University of Maryland, Baltimore County,
1000 Hilltop Circle, Baltimore, MD 21250, USA.
niyatic1@umbc.edu

March 15, 2011

1 Abstract

The major contributions of this work include a set of biographical feature extractors brought together by a probabilistic graphical model. The graphical model acts as a communication media between different feature extractors, together resulting in a text descriptor to describe triage images of disaster victims. The model is built using domain information gathered from data and literature.

2 Motivation

Earthquakes, hurricanes, terrorist attacks, and other such events can cause tremendous harm to people and infrastructure, often leaving those outside the areas impacted with little or no information on the state of their friends and relatives who may have been affected. Our goal is to automatically process images of patients taken as part of the intake process at emergency medical care centers to extract searchable, text-based descriptors of patients that can be accessed remotely (e.g., via the web) to facilitate identification of victims. Lost Person Finder [1] is a research project which aims at providing immediate post disaster communication information about disaster victims. As a part of the project, triage images of the victims are gathered and the aim is to extract a text descriptor that describes the person in the image. Achieving this goal requires finding the person in the image, and in particular their face, to drive the process of extracting features such as hair color and style, clothing color and style, the presence or absence of glasses, and other similar features.

Face detection for such non-standard images is a difficult task and was the motivation behind starting work in this domain. As we dealt with each feature detector one at a time, we realized that the different biographical features are dependent on each other. The target is to exploit this inter-dependency between individual feature detectors and improve the efficiency of the complete system.

3 Related Work

Extraction of biographical features has been studied to a limited extent by different computer vision groups. There is individual work in age detection [7], gender identification [5], ethnicity recognition [8], eyeglasses identification [6], hair detection and hair style recognition [10]. Most of this work is limited and addressed to a particular set of images and tends to do poorly with disaster victim images. At the same time, there is no particular work that brings together all these features and tries to leverage relationships between them. The computer vision community has done some work in topics such as scene completion where information about different parts of the image is available, but this is a completely different problem than the one we are dealing with.

4 Approach

Feature extraction as introduced above needs to be preceded by first locating the person in the image, particularly the face. Face detection has been extensively studied for standard template-like face images, but for disaster victim images standard techniques

do not perform well. In this work face detection [3] is done as a pre-processing step. Individual feature extractors are studied and then the extractors link with each other into a probabilistic graphical model. The individual feature detectors give values for some biographical feature as a output. For example, gender detector output is about whether the person in the image is a male or a female. Inference algorithm applied over this model help come to an agreement between the different extractors and, on convergence of this algorithm, a text descriptor is built using output from different extractors, which gives identification information about the person in the image. Figure 1 shows an overview of this approach.

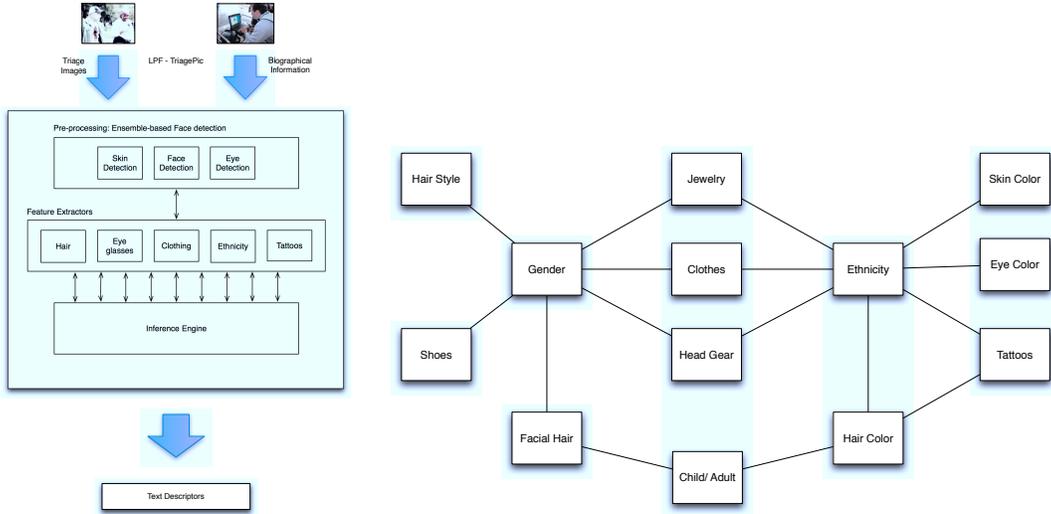


Figure 1: System architecture and Markov Net

As stated earlier, we initially developed an ensemble-based face detection algorithm that would perform well for images of mass disaster victims. This algorithm is a combination of skin detection, face detection and a template-based eye detection algorithm. Our dataset is a set of triage images of disaster victims, gathered during mass disaster drills conducted for Lost Person Finder. This algorithm gave a F-Score of 0.8 as against standard algorithms like Viola-Jones [9], which gave a performance of 0.45. Experiments for each stage of this algorithm, i.e. skin detection, face detection and face detection with a combination of skin and eye detection, show that the ensemble performs best. The next step is addressing the various different feature extractors. So far, we have done work in eyeglasses detection, hair region detection using a combination of color and texture analysis, hair color and skin color identification and preliminary work in ethnicity identification. Preliminary experiments show these developed approaches perform better than existing algorithms for the given set of images.

The main aim of this work is to bring these different extractors together. We designed an ad-hoc Markov net using information available in the literature. That is, information such as gender and facial hair tend to be related to each other and so do age and hair color. Figure 1 shows a diagram of this ad-hoc graph. We further developed a factor graph for this net. We are also working on gathering data to validate the different edges in the graph and to help define the domains of different nodes in a more complete manner using person datasets [4] [2] and the Amazon mechanical turk.

5 Conclusion and Future work

In this work, we present an approach to design a probabilistic graphical model to build a text descriptor that would help identify a person, given the image of the person. We present the problem as a two phase approach, one that involves computer vision and image processing techniques to extract required features and the second that builds and learns using a probabilistic graphical model that uses relationships between the different biographical features.

Having completed the initial graphical design, we are currently working on gathering data to validate the relationships. We are also working on defining the potential functions that communicate between the adjoining nodes and indicate where the nodes agree with each other or not. Our principal approach is to use energy minimization where higher the value of the potential function lower is the agreement between the nodes. This information will help us validate the complete text descriptor, hence making it more useful for real life applications.

References

- [1] Lost person finder. <http://archive.nlm.nih.gov/proj/lpf.php>.
- [2] Tamara L. Berg, Alexander C. Berg, Jaety Edwards, and David A. Forsyth. Whos in the picture. In *NIPS*, 2004.
- [3] Niyati Chhaya and Tim Oates. Robust face detection for patient triage images. In *VISAPP 2011- Proceedings of the Sixth International Conference on Computer Vision Theory and Applications, to appear*. INSTICC Press, 2011.
- [4] A. Gallagher and T. Chen. Clothing cosegmentation for recognizing people. In *Proceeding of CVPR*, 2008.
- [5] Arnulf B. A. Graf and Felix A. Wichmann. Gender classification of human faces. In *Biologically Motivated Computer Vision 2002, LNCS 2525*, pages 491–501, 2002.
- [6] Xiaoyi Jiang, Michael Binkert, and Bernard Achermann. Towards detection of glasses in facial images. *Pattern Analysis Applications*, 3:9–18, 2000.
- [7] Young H. Kwon and Niels da Vitoria Lobo. Age classification from facial images. *Computer Vision and Image Understanding*, 74:1–21, April 1999.
- [8] Xiaoguang Lu and Anil K. Jain. Ethnicity identification from face images. In *SPIE International Symposium on Defense and Security : Biometric Technology for Human Identification*, pages 114–123, 2004.
- [9] Paul Viola and Michael Jones. Robust real-time object detection. In *International Journal of Computer Vision*, 2001.
- [10] Yaser Yacoob and Larry S. Davis. Detection and analysis of hair. *IEEE Transactions Pattern Analysis Machine Intelligence*, 28(7):1164–1169, 2006.