

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

Niyati Chhaya and Tim Oates

University of Maryland, Baltimore County
1000 Hilltop Circle, Baltimore, Maryland, USA

{niyatic1, oates}@umbc.edu

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 (LPF,) 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 our 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, cluttered 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.

In this work, individual feature extractors are studied and then the extractors link with each other into a probabilistic graphical model. The feature detectors give values for a particular soft biometric feature as a output. For example, gender detector output is about whether the person in the image is a

male or a female. Inference algorithms 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.

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. An Amazon mechanical turk survey is used to understand relationships between the individual nodes. The annotations aid in understand how people describe other people. The annotators were asked to describe the person in the given image and mark a few soft biometric features as required for person identification. This data will be used to design the text descriptor according to its utility.

We use a message passing-based inference to establish agreement between individual nodes. Potential functions defined are based on the confidence of feature extractor, validity of feature extractor values, agreement between different nodes and a temporal penalty term to force early convergence of the inference algorithm. In our current experiments, we present the behavior of potential values with different situations in the node agreements. The inference and the graphical model follow an energy minimization approach.

The final aim of this work is to establish interaction between the graphical model, the inference algorithm, and the feature extractors. This would help in making the text descriptor robust and aid in searching a particular person from a large dataset.

References

- Niyati Chhaya. 2011. Joint inference for extracting text descriptors from triage images of mass disaster victims.
- Lost person finder. <http://archive.nlm.nih.gov/proj/lpf.php>.