Enhanced IR face recognition using Bayesian unsupervised algorithm

Tarek Elguebaly and Niza Bouguila

An algorithm based on the generalized Gaussian-mixture model offers robustness to noise and flexibility in modeling a wide range of facial appearances.

In recent years, face recognition has become one of the most rapidly growing research areas in computer vision because of its wide range of potential applications related to security and safety. Face recognition is the most natural, intuitive way to identify individuals compared to several other biometric authentication methods such as fingerprints, iris patterns, and voiceprint that generally rely on cooperation of the participants. Given an unknown facial input image, the system identifies the person. This process is generally based on three main tasks: face detection, feature extraction, and face identification. Face detection is used to distinguish face-like from other objects. To decrease the dimensionality of face images, they are usually represented in terms of low-level feature vectors in lower-dimensional feature space (i.e., face signature) for recognition. The face-identification task identifies the input person’s face by searching a database of known individuals.

Despite the variety of approaches and tools available, face recognition is still a challenging problem because of both the variations in appearance that a given face may have and numerous image-acquisition problems such as noise, video-camera distortion, and image resolution. This is especially the case for face recognition in the visible spectrum, since performance is sensitive to variations in illumination conditions, even more than to changes in face identity. Thermal IR offers a promising alternative for handling variations in face appearance caused by illumination changes, facial expressions, and face poses. Thus, IR-based algorithms have the potential to provide simpler and more robust solutions, improving performance in uncontrolled environments and combating deliberate attempts to obscure identity.

Although illumination and facial expression significantly change the visual appearance of a face, its thermal characteristics remain nearly invariant (see Figure 1). Several approaches have been proposed to analyze and recognize faces at IR wavelengths. They can be divided into appearance- and feature-based methods. While the former focus on a face’s global properties, the latter explore the facial features’ (e.g., eyes or mouth) statistical and geometric properties. Many of these approaches, however, assume that the extracted IR face features are Gaussian, which is not generally an appropriate assumption. We have proposed an appearance-based approach using unsupervised Bayesian learning of finite, generalized Gaussian-mixture models.

We considered face recognition as an image-classification problem by trying to determine to which person the image belongs. This can be divided into two main tasks, including feature extraction and similarity measurements. Feature extraction allowed us to generate a set of features known as image signatures capable of representing the visual content of each image in a lower-dimensional space. We used similarity measurements to assign each image to the right group. For the feature-extraction step, we employed both edge-orientation histograms (which provide spatial information)
and co-occurrence matrices, which capture the local spatial relationships between gray levels.\textsuperscript{12} Edge-orientation histograms can be very successful at distinguishing between faces (see Figure 2, where it is quite clear that the first two images have very similar edge-orientation histograms compared to the third image). For similarity measurements, we used mixture models because they can model complex datasets.

In most image-processing and computer-vision applications, a Gaussian density is applied for data modeling. However, most data generated from computer-vision applications is non-Gaussian. Many studies have demonstrated that the generalized Gaussian distribution (GGD) can be a good alternative thanks to its shape flexibility, which allows modeling of a large number of non-Gaussian signals.\textsuperscript{13–16} The GGD is defined by three parameters that represent the mean, the inverse scale, and the shape parameter of the $d$-dimensional GGD. The complete data likelihood for this case is

$$P(X|\theta) = \sum_{j=1}^{M} P(X|\theta_j) p_j,$$

where $p_j$ are the mixing proportions, which must be positive and sum to one. $\theta = (\mu_1, \ldots, \mu_M, \alpha_1, \ldots, \alpha_M, \beta_1, \ldots, \beta_M, p_1, \ldots, p_M)$ is the set of parameters of the mixture with $M$ classes, where $\mu = (\mu_1, \ldots, \mu_d)$, $\alpha = (\alpha_1, \ldots, \alpha_d)$, and $\beta = (\beta_1, \ldots, \beta_d)$ are the mean, the inverse scale, and the shape parameter of the

Approaches for estimation of finite-mixture-model parameters can be arranged into two categories, i.e., deterministic and Bayesian methods. In the former, parameters are taken as fixed and unknown, and inference is founded on the likelihood of the data. On the other hand, Bayesian methods are related to Bayesian theory, which means that they allow for probability statements to be made directly about the unknown parameters of the mixture, while taking into consideration prior or expert opinion. The goal is to get the posterior distribution, $\pi(\theta|X, Z)$, by combining the prior information about the parameters, $\pi(\theta)$, with the observed value or realization of the complete data set, $P(X, Z|\theta)$, which is derived from Bayes formula,

$$\pi(\theta|X, Z) = \frac{\pi(\theta) P(X, Z|\theta)}{\int_\theta \pi(\theta) P(X, Z|\theta) d\theta} \propto \pi(\theta) P(X, Z|\theta),$$

where $(X, Z)$ is the complete data set. With $\pi(\theta|X, Z)$ in hand, we can simulate our model parameters, $\theta$, rather than computing them. We based our learning algorithm on the Monte Carlo simulation technique of Gibbs sampling, mixed with a Metropolis-Hasting step.\textsuperscript{18} We used the integrated or marginal likelihood using the Laplace-Metropolis estimator to obtain $M$.\textsuperscript{19}

We applied our approach to images from the Iris thermal face database, which is a subset of the ‘Object tracking and classification beyond the visible spectrum’ database.\textsuperscript{9} First, we used 1320 images of 15 persons not wearing glasses. Knowing that in IR imaging thermal radiation cannot transmit through glasses because the latter severely attenuate electromagnetic-wave radiation beyond 2.4mm, we decided to investigate whether our algorithm could identify persons wearing glasses, and so we added 880 images of eight persons with glasses. For both experiments, we used 11 images for each person as training set and

Continued on next page
the rest for testing. This gave us 165 and 1155 images in the first data set for training and testing, respectively. The second data set contained 253 and 1947 images for training and testing, respectively. To validate our Bayesian algorithm, we compared it with an expectation-maximization algorithm. We also compared it to six other methods (see Table 1). Bayesian learning outperformed all other methods because of its ability to incorporate prior information during learning and modeling of classes.

In summary, our new algorithm for IR face recognition based on Bayesian learning of finite, generalized Gaussian-mixture models is accurate and effective. We chose the generalized Gaussian distribution for its flexibility and because of the fact that IR image statistics are generally non-Gaussian. In contrast with the expectation-maximization algorithm, which tries to estimate a single ‘best’ model that is not generally realistic, Bayesian approaches take into account the fact that the data may suggest many ‘good’ models and consider the average result computed over several models. We are now developing a variational framework for the learning of the generalized Gaussian-mixture model and its application to other challenging problems such as fusion of visual and thermal spectral modalities.

Table 1. Accuracies for seven different methods. BGGM, BGMM, EMGGM, EMGMM: Bayesian generalized Gaussian, Bayesian Gaussian mixture, expectation-maximization generalized Gaussian, and Gaussian mixture models learned with expectation maximization. PCA: Principal-component analysis. HICA: Holistic independent-component analysis (ICA) with cosine distance as implemented by fixed-point algorithm (FastICA).20 LICA: Localized ICA.

<table>
<thead>
<tr>
<th>Method</th>
<th>Data set 1</th>
<th>Data set 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>BGGM</td>
<td>96.02%</td>
<td>95.33%</td>
</tr>
<tr>
<td>EMGGM</td>
<td>94.20%</td>
<td>92.40%</td>
</tr>
<tr>
<td>BGMM</td>
<td>86.67%</td>
<td>82.54%</td>
</tr>
<tr>
<td>EMGMM</td>
<td>85.89%</td>
<td>82.18%</td>
</tr>
<tr>
<td>PCA</td>
<td>95.58%</td>
<td>94.35%</td>
</tr>
<tr>
<td>HICA</td>
<td>95.32%</td>
<td>93.99%</td>
</tr>
<tr>
<td>LICA</td>
<td>94.46%</td>
<td>92.81%</td>
</tr>
</tbody>
</table>

References

Author Information
Tarek Elguebaly and Niza Bouguila
Concordia Institute for Information Systems Engineering
Concordia University
Montreal, Canada

Tarek Elguebaly received a bachelor’s degree in electronics and communication engineering from Cairo University (Egypt) in 2006 and a master’s from Concordia University in 2009. His research focuses on machine learning and computer vision.

Nizar Bouguila received an engineering degree from the University of Tunis (Tunisia) in 2000 and a computer science MSc and PhD from Sherbrooke University in 2002 and 2006, respectively. He is currently an associate professor.