Novel composite operator for cross-spectral face recognition

Zhicheng Cao and Natalia Schmid

The Composite Multi-Lobe Descriptor shows promising results and outperforms other popular operators.

Face recognition algorithms are fully automated methods for identifying and verifying individuals, based on images of their faces. Over the past few decades many major advances have been made in this area. As such, applications of face recognition have expanded far beyond identity authentication, surveillance, access control, and information security in commercial, military, and law enforcement scenarios. Despite the important breakthroughs in the field, however, several challenges have arisen and many issues remain unsolved. For instance, the introduction of novel imaging modalities has led to new research opportunities and has created many new research problems. For example, short-wave IR (SWIR) cameras have recently become commercially available. With these cameras it is possible to ‘see’ at night through fog, rain, and in harsh environments. In addition, sharp face images can be generated at long standoff distances with these SWIR cameras. One of the most important outstanding issues in cross-spectral face recognition studies is therefore the development of new algorithms that are capable of processing and matching SWIR, near-IR (NIR), mid-wave IR, and long-wave IR face images with a gallery of visible light images (see Figure 1).

Figure 1. Sample images at different wavelengths. (a) Visible light. (b–d) Short-wave IR (SWIR), taken at standoff distances of 1.5, 50, and 106m, respectively. (e–g) Near-IR, taken at standoff distances of 1.5, 50, and 106m, respectively.

We have introduced a new composite operator—known as the Composite Multi-Lobe Descriptor (CMLD)—for cross-spectral face recognition, with which we take a local operator-based approach. We have demonstrated the performance of the CMLD on datasets of SWIR and NIR face images, which we have matched to a gallery of visible light images. In our operator, we combine Gabor filters, local binary patterns (LBPs), generalized LBP (GLBPs), and the Weber local descriptor (WLD), and we modify them into a multi-lobe function with smoothed neighborhood. We use the LBP, GLBP, and WLD to encode the intensity and orientation information in the patterns that are formed by extracted edges. In addition, our use of multi-lobe functions with smoothed neighborhoods makes our proposed operator robust against noise and poor-quality images. We map our encoded images into a histogram representation, and we then cross-match them by applying a symmetric Kullback-Leibler distance.

Continued on next page
It has recently been shown that a two-step feature extraction procedure—with Gabor filters followed by multi-lobe ordinal measures—yields an outstanding performance for recognition when it is applied to visible face images.\textsuperscript{12} We were inspired by this result to develop our new CMLD operator.\textsuperscript{13} The first multi-lobe operator we use transforms the original LBP into a multi-lobe LBP (MLLBP). In this process, discrete values of LBP are replaced with smooth Gaussian functions that have multiple lobes (with alternating positive and negative signs). Each resulting Gaussian lobe carries a balanced weight. To encode a face image, we apply the MLLBP to a preprocessed image, and then to a unit step function and uniform pattern mapping. For our multi-lobe version of a GLBP (MLGLBP), we use the same principle as for the MLLBP. The last step in the process, however, is different. For the MLGLBP, we use a threshold operator followed by uniform pattern mapping. Our multi-lobe WLD (MLWLD) is also a modification of the original WLD operator. We achieve this modification by using a concept that is similar to the one we used for LBP, but which also involves normalization of the intensity difference between a central pixel and its neighbors. In the final building steps for our MLWLD, we apply an inverse tangent function and an L-level quantizer (that maps continuous values within a given range to several—L—discrete values).

Before we apply any operator to an image, we perform a few basic preprocessing steps. These include face image alignment, cropping, and normalization. For face image alignment, we select the center of the eyes and project these to fixed positions using geometric transformations (i.e., rotation, scaling, and translation). We then crop and normalize the aligned face images. For normalization, we make a color-to-grayscale conversion for the color images, and a log normalization of the image intensity for SWIR and NIR images. The results from our encoding procedures—on the sample images shown in Figure 1(a)—are illustrated in Figure 2.

We have tested the performance of our new operator on the Pre-Tactical Imager for Night/Day Extended-Range Surveillance (TINDERS) and TINDERS datasets. These datasets were collected by the Advanced Technologies Group of the West Virginia High Technology Consortium Foundation.\textsuperscript{14} The Pre-TINDERS data was obtained at a short standoff distance (1.5m) between the subject and the camera. For the TINDERS data, the standoff distances were longer (50 and 106m). Both the datasets are composed of 48 subject images in frontal view, and each subject is represented by four or five images that were collected in two separate sessions for each visible, NIR, and SWIR spectral band (as illustrated in Figure 1). The total number of images in the Pre-TINDERS and TINDERS datasets are therefore 576 and 1255, respectively. To ensure a fair comparison between the two datasets, we cropped and normalized them so that they were the same size. We compare the recognition performance of our

\begin{figure}[h]
\centering
\includegraphics[width=\linewidth]{figure2.png}
\caption{Results of encoding with different operators. (a) and (b) Magnitude and phase of a Gabor filter response. (c–e) Images after the application of a multi-lobe local binary pattern (LBP), a multi-lobe generalized LBP, and a multi-lobe Weber local descriptor, respectively.}
\end{figure}

\begin{figure}[h]
\centering
\includegraphics[width=\linewidth]{figure3.png}
\caption{The performance of the composite multi-lobe descriptor (CMLD) for matching SWIR and visible face images, at a standoff distance of 1.5m. These receiver operating characteristic curves show that the CMLD operator compares well with other popular operators, such as LBP, histogram of oriented gradients (HOG), a Gabor filter followed by LBP, and a Gabor ordinal measure (GOM). FAR: False acceptance rate. GAR: Genuine acceptance rate.}
\end{figure
CMLD operator and several other operators in Figure 3, and it is clear that our CMLD outperforms the other four operators. SWIR images at long standoff distances (i.e., 50 and 106 m) experience some loss of quality because of air turbulence, insufficient illumination, and optical effects during data acquisition. All the operators therefore experience a drop in performance level at these longer standoff distances. Our CMLD operator, however, remains the leading option in all cases. We have also obtained similar results by matching NIR face images to a gallery of visible face images, for both short and long standoff distances.

Our new CMLD operator can be used for facial feature extraction and therefore can be used to address outstanding problems in the field of cross-spectral face recognition. We have successfully demonstrated that the operator can be used to match NIR and SWIR probes against a gallery of visible light face images. We have also analyzed the performance of the CMLD and have shown that it substantially outperforms other commonly used operators. We now plan to extend the idea of multi-lobe smoothing functions to other operators. We also hope to investigate the performance of our CMLD operator for cross-spectral periorificial recognition purposes, and to use our operator for other biometric problems (e.g., face detection at IR wavelengths).

We thank Brian Lemoff from the West Virginia High Technology Consortium Foundation for providing the Pre-TINDERS and TINDERS datasets that were employed in this study.

Author Information

Zhicheng Cao and Natalia Schmid
Lane Department of Computer Science and Electrical Engineering
West Virginia University
Morgantown, WV

Zhicheng Cao is a PhD student. He obtained his BS and MS in biomedical engineering from Xi’an Jiaotong University, China. His research interests include biometrics, computer vision, pattern recognition, and image processing with a focus on cross-spectral recognition of the face, periocular regions, and partial faces.

Natalia Schmid received her PhD in engineering from the Russian Academy of Sciences in Moscow, Russia, and her DSc in electrical engineering from Washington University, MO. She is now an associate professor. Her current research interests include detection and estimation, learning theory and information theory with application to distributed sensor networks, biometric and forensic systems, and signal processing for radio astronomy applications.

References