

Face Recognition with Disparity Corrected Gabor Phase Differences

Manuel Günther¹, Dennis Haufe², and Rolf P. Würtz²

¹ Idiap Research Institute, Martigny, Switzerland

² Institut für Neuroinformatik, Ruhr Universität, Bochum, Germany

Abstract. We analyze the relative relevance of Gabor amplitudes and phases for face recognition. We propose an algorithm to reliably estimate offset point disparities from phase differences and show that disparity-corrected Gabor phase differences are well suited for face recognition in difficult lighting conditions. The method reaches 74.8% recognition rate on the Lighting set of the CAS-PEAL database and 35.7% verification rate on experiment 2.4 of the FRGC database.

1 Introduction

Gabor wavelets are widely used for face recognition [10,2,11,1] and as models for processing in the primary visual cortex. The complex-valued Gabor responses are split up into an amplitude and a phase, which is called *Gabor phase*. The amplitudes are models for the responses of complex cells in the visual system. Recently, [12] achieved good face recognition from coarsely discretized Gabor phases (extended local Gabor binary pattern), especially under strong illumination conditions.

Gabor wavelet responses at single locations of facial images are collected into Gabor jets [10], which are extracted at several offset positions and assembled into a Gabor graph \mathcal{G} . Often, face graphs with nodes at facial landmark positions are utilized [10,2], but in the present paper we use rectangular grid graphs, which have shown to be at least as expedient [1,8].

The texture contents of Gabor graphs can be displayed by reconstruction [2]. Exemplary reconstructions of grid graphs are shown in Figure 1, where Figures 1(a) and 1(b) show the aligned images \mathcal{I}_A and \mathcal{I}_B , while Figures 1(c) and 1(d) display the images reconstructed from the Gabor graphs \mathcal{G}_A and \mathcal{G}_B . To show the significance of Gabor phases for face recognition, absolute values and phase values from graphs \mathcal{G}_A and \mathcal{G}_B are merged into $\mathcal{G}_{A,B}$ and $\mathcal{G}_{B,A}$, where each Gabor jet holds amplitudes of \mathcal{G}_A and phases of \mathcal{G}_B or vice versa. Visual inspection of both reconstructions reveals that they resemble strongly the face whose phases have been used. This can be interpreted such that recognition is coupled to Gabor phases rather than amplitudes.

In this paper we first review the basics of Gabor graphs, then define different similarity functions with and without disparity estimation. Then we test them on two standard datasets, find good recognition rates using disparity, which are finally discussed.

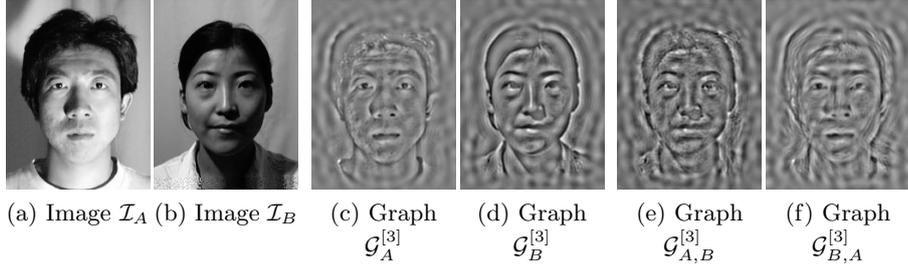


Fig. 1. Reconstruction from graphs with exchanged Gabor phases: (a) and (b) show the preprocessed images, (c) and (d) show reconstructions of the texture information included in the according grid graphs of type $\mathcal{G}^{[3]}$ (cf. Section 4). (e) and (f) display reconstructed graphs containing the absolute Gabor wavelet responses from (c) and the phase values from (d), and vice versa.

2 Gabor jets and their similarities

Gabor wavelets $\psi_{\underline{k}}(\underline{x})$ are two-dimensional complex-valued image filters [6]:

$$\psi_{\underline{k}}(\underline{x}) = \frac{k^2}{\sigma^2} \exp\left(-\frac{k^2 \underline{x}^2}{2\sigma^2}\right) \left[\exp\left(i \underline{k}^T \underline{x}\right) - \exp\left(-\frac{\sigma^2}{2}\right) \right]. \quad (1)$$

Commonly, the discrete Gabor wavelet family is generated by defining 40 different parameter vectors \underline{k}_j [10]:

$$\underline{k}_j = \underline{k}_{\mu,\nu} = \begin{pmatrix} k_\mu \cos \varphi_\nu \\ k_\mu \sin \varphi_\nu \end{pmatrix}; \quad k_\mu = 2^{-\frac{2+\mu}{2}} \pi; \quad \varphi_\nu = \frac{\nu \pi}{8}; \quad j = 8\mu + \nu \quad (2)$$

in five scales ($\mu = 0, \dots, 4$) and eight orientations ($\nu = 0, \dots, 7$).

A *Gabor jet* \mathcal{J} is the collection of these 40 responses of these Gabor wavelets at a given image position, generated by the convolution of \mathcal{I} with all Gabor wavelets:

$$(\mathcal{J}(\underline{x}))_j = (\mathcal{I} * \psi_{\underline{k}_j})(\underline{x}). \quad (3)$$

The complex-valued elements $(\mathcal{J})_j$ of the Gabor wavelets are used in normalized polar form: $\|\mathcal{J}\| \cdot a_j \cdot \exp(i \phi_j)$, with *phases* ϕ_j and *amplitudes* a_j normalized by the length of the Gabor jet [5].

A Gabor graph consists of a set of nodes, which are located at different positions \underline{x}_n and connected by a set of edges. Both number and position of nodes and graph topology are, in principle, arbitrary. In this paper we restrict the positions to rectangular grids $\underline{x}_{\alpha\beta} = x_{00} + \alpha(0, \Delta)^T + \beta(\Delta, 0)^T$, $\alpha \in \{0, \dots, \alpha_{\max}\}$, $\beta \in \{0, \dots, \beta_{\max}\}$. Unlike in [6], all similarity functions studied here are independent of the edges, and therefore the edge set does not influence the results. For visualization, edges connect the four nearest neighbors. To assure that corresponding facial features are compared, images are aligned to the hand-labeled eye positions provided with the test databases, and one node is placed at each eye center

which yields the value of Δ . To study the influence of both grid sizes and image resolution we use grids $\mathcal{G}^{[1]}$ with 7×5 nodes, $\mathcal{G}^{[2]}$ with 10×8 nodes, Grid $\mathcal{G}^{[3]}$ with 23×16 nodes on 168×224 pixel images, and $\mathcal{G}^{[4]}$ with 23×16 nodes on 300×400 pixel images.

To estimate the identity of a probe image, it is compared with several gallery images and assigned the identity of the most similar gallery image. Image comparison is traced back to the comparison of the two Gabor graphs extracted from these images. Two graphs are compared by computing similarities of Gabor jets at corresponding grid positions. Due to head rotation, facial expressions, or image alignment errors, the offset positions of the two corresponding Gabor jets might be incorrect. In the *elastic bunch graph matching* (EBGM) approach [10], the correspondence problem is solved locally by computing offset position corrections, so-called *disparities*. Applying these disparities leads to distorted graphs [10], where only integral offset position errors are corrected.

To use Gabor phase differences reasonably, sub-pixel accuracy is required since disparities of half a pixel length can change Gabor phases by up to $\pi/4$. For estimating the disparities in sub-pixel accuracy, a well-known technique [9,10] processing Gabor jets is extended to estimate larger disparity vectors reliably. Instead of moving node positions, these disparities are used directly to correct the Gabor phase difference in the face recognition process. Furthermore, we propose a simple method to incorporate absolute and phase values of Gabor wavelet responses. We show that this method works well on images with controlled frontal illumination, and outperforms current face recognition algorithms on images with strong non-frontal or uncontrolled illumination.

Jets are compared by similarity functions, some of them are based on amplitudes only (see [3]). The following functions were already successfully employed [10,2] to solve the correspondence problem and/or calculate global similarities for face recognition (S_A , S_C , and S_D stand for amplitude, Canberra, and disparity similarity, respectively):

$$S_A(\mathcal{J}, \mathcal{J}') = \sum_{j=0}^{39} a_j a'_j, \quad (4)$$

$$S_C(\mathcal{J}, \mathcal{J}') = 1 - \frac{1}{40} \sum_{j=0}^{39} \frac{|a_j - a'_j|}{a_j + a'_j}, \quad (5)$$

$$S_D(\mathcal{J}, \mathcal{J}') = \sum_{j=0}^{39} a_j a'_j \cos\left(\phi - \phi' - \underline{k}_j^T \underline{d}\right). \quad (6)$$

The *disparity vector* $\underline{d} = (d_h \ d_v)^T$ accounts for the correction between the offset positions of \mathcal{J} and \mathcal{J}' . Given a Gabor jet \mathcal{J} at a certain position in a reference image, the disparity vector of Gabor jet \mathcal{J}' in the probe image points to the location where the Gabor jet most similar to \mathcal{J} is to be expected. In the case of S_D , the disparity vector \underline{d} is estimated by solving eq. (10) below by a new coarse-to fine strategy. The respective similarity functions are (S_N and S_n for two types of “new” similarity):

$$S_N(\mathcal{J}, \mathcal{J}') = \frac{1}{40} \sum_{j=0}^{39} a_j a'_j \cos\left(\phi_j - \phi'_j - \underline{k}_j^T \underline{d}_0\right), \quad (7)$$

$$S_n(\mathcal{J}, \mathcal{J}') = \frac{1}{40} \sum_{j=0}^{39} \cos\left(\phi_j - \phi'_j - \underline{k}_j^T \underline{d}_0\right). \quad (8)$$

Finally, we use the following alternative combination:

$$S_{n+C}(\mathcal{J}, \mathcal{J}') = S_n(\mathcal{J}, \mathcal{J}') + S_C(\mathcal{J}, \mathcal{J}'). \quad (9)$$

The similarity of two grid graphs is computed as the average similarity of the Gabor jets at the grid positions.

3 Disparity estimation

In [10] the Taylor expansion $\cos(x) \approx 1 - \frac{1}{2}x^2$ was used to approximate the similarity S_D from eq. (6), and the optimal disparity is calculated by setting the gradient of the resulting function to zero. This leads to [9,2] $\underline{d} = \mathbf{\Gamma}^{-1} \mathbf{\Phi}$, with

$$\begin{aligned} \mathbf{\Gamma}_{h,v} &= \sum_{j=0}^{39} a_j a'_j k_{j;h} k_{j;v} \\ \mathbf{\Phi}_h &= \sum_{j=0}^{39} a_j a'_j k_{j;h} (\phi_j - \phi'_j) \end{aligned} \quad (10)$$

The resulting disparity vector \underline{d} is only useful in a very limited range around the offset point (see Figure 2(a)) because complete phase cycles are not addressed correctly in the Taylor expansion. In this paper, we employ a course-fine strategy to estimate complete 2π phase cycles. We replace the linear offset in eq. (10) by adding an integral multiple n_j of the wavelength of the respective wavelets, i. e., we approximate $\cos(x) \approx 1 - \frac{1}{2}(x - n2\pi)^2$ for $x \approx n2\pi$. This leads to the following algorithm: At the lowest frequency scale $\mu = 4$, all n_j are set to zero. The disparities: $\underline{d}_\mu = \mathbf{\Gamma}_\mu^{-1} \mathbf{\Phi}_\mu$ are calculated using:

$$\begin{aligned} \mathbf{\Gamma}_{\mu;h,v} &= \sum_{j=8\mu}^{39} a_j a'_j k_{j;h} k_{j;v} \\ \mathbf{\Phi}_{\mu;h} &= \sum_{j=8\mu}^{39} a_j a'_j k_{j;h} (\phi_j - \phi'_j - n_j 2\pi). \end{aligned} \quad (11)$$

For the following scales $\mu = 3, \dots, 0$, the n_j values are computed using the estimated disparity from the lower frequency levels:

$$\forall j \in \{8\mu, \dots, 8(\mu+1) - 1\} \quad : \quad n_j = \left\lfloor \frac{\phi_j - \phi'_j - \underline{k}_j^T \underline{d}_{\mu+1}}{2\pi} \right\rfloor \quad (12)$$

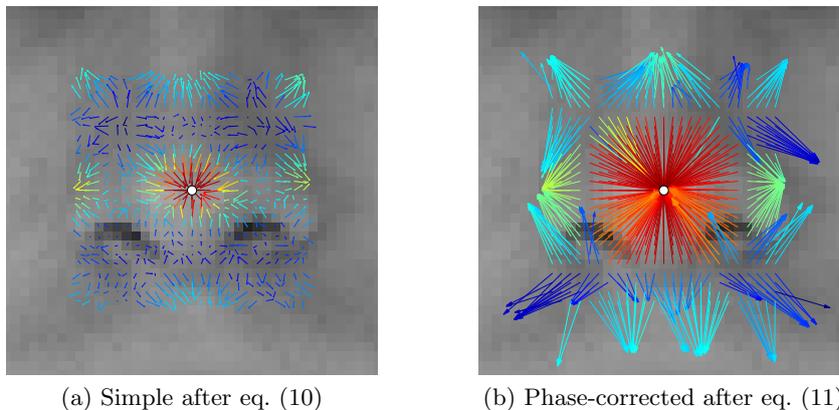


Fig. 2. Disparity estimation: Each arrow shows the disparity \underline{d} (a) or \underline{d}_0 (b) from the Gabor jet at its offset point and the reference Gabor jet at the nose-tip (marked with a white dot). The color codes the similarity value, ranging from blue (low) to red (high).

where $\lfloor \cdot \rfloor$ denotes the integral rounding operator. The final disparity is \underline{d}_0 , which corresponds much better to the true offset (Figure 2(b)).

4 Experiments

The face recognition or verification results of the different similarity functions are computed for two image databases: the CAS-PEAL database [1] and the FRGC database in the version ver2.0 [7]. Both provide hand-labeled eye positions, experimental setups and according baseline results, which are compared to the results of our recognition system.

4.1 CAS-PEAL

The CAS-PEAL [1] images are partitioned into a gallery of 1040 images with ambient illumination and neutral facial expression, and different probe sets. We here process the probe sets *Expression* and *Lighting*. For each probe set, the four different grid types and several similarity functions are tested. In case of grid size $\mathcal{G}^{[4]}$, the images of the original resolution of 360×480 pixels are partially scaled up before cutting out the face, possibly leaving scaling artifacts. The best results are summarized in Table 1.

The *Expression* probe set contains 1570 images with ambient illumination, showing one of five facial expressions. Generally, the smaller graphs performed worse than $\mathcal{G}^{[3]}$ and $\mathcal{G}^{[4]}$. The phase-corrected disparity similarity functions S_N and S_n reach rates about 95%. The combination $S_n + C$ is slightly worse than S_C alone. The performance of $\mathcal{G}^{[4]}$ is inferior to $\mathcal{G}^{[3]}$, replicating the findings of [4,5].

Algorithm	<i>Expression</i>	<i>Lighting</i>
PCA	53.7%	8.2%
PCA+LDA	71.3%	21.8%
GPCA+LDA	90.6%	44.8%
LGBPHS	95%	51%
S_C on $\mathcal{G}^{[3]}$	98.9%	59.6%
S_n on $\mathcal{G}^{[3]}$	97.2%	59.7%
S_{n+C} on $\mathcal{G}^{[3]}$	98.3%	67.9%
S_{n+C} on $\mathcal{G}^{[4]}$	96.8%	74.8%

Table 1. Result comparison on CAS-PEAL: Baseline results from [1] and the best results of our recognition experiments. Column maxima are in boldface.

The *Lighting* probe set consists of 2243 images with neutral expression, but strong fluorescent illumination from fifteen different directions and one frontal incandescent illumination. Recognizing these probe images by comparing them to frontally illuminated gallery faces is hard. The results obtained using $\mathcal{G}^{[3]}$ and $\mathcal{G}^{[4]}$ again outperform the low resolution grids $\mathcal{G}^{[1]}$ – $\mathcal{G}^{[2]}$. We [4] showed that higher image resolutions and higher frequency Gabor wavelets lead to the same results, indicating that the phases of higher frequencies are more valuable, at least when dealing with strong non-frontal illumination.

The CAS-PEAL database [1] provides the recognition results for eigenface-based algorithms like PCA or PCA+LDA and more recent Gabor-wavelet-based algorithms like GPCA+LDA and LGBPHS. In Table 1, they are compared with the results reported in this paper. Recognition rates are directly comparable since identical gallery and probe sets are taken, and also [1] aligned the facial images according to the same hand-labeled eye positions. The Gabor-wavelet-based GPCA+LDA, which also requires training data, and the non-trained LGBPHS show high recognition rates on the *Expression* and *Lighting* sets. The results reported in this paper, which use untrained image comparisons, are a little higher on the *Expression* subset, and outperforming the *Lighting* subset by far. The combination S_{n+C} on $\mathcal{G}^{[4]}$ gives half the error rate in the *Lighting* experiments.

4.2 FRGC

The *Face Recognition Grand Challenge* (FRGC) database in version 2.0 [7] consists of 36818 facial images of 466 persons taken under controlled or uncontrolled lighting with some expression. They are divided into a training set of 12776 images of 222 identities and a target set of 16028 controlled and a query set of 8014 uncontrolled images of all 466 people. Here, we perform experiment 2.1, comparing target with target images, and experiment 2.4, matching query images with uncontrolled illumination against the target images with controlled illumination.

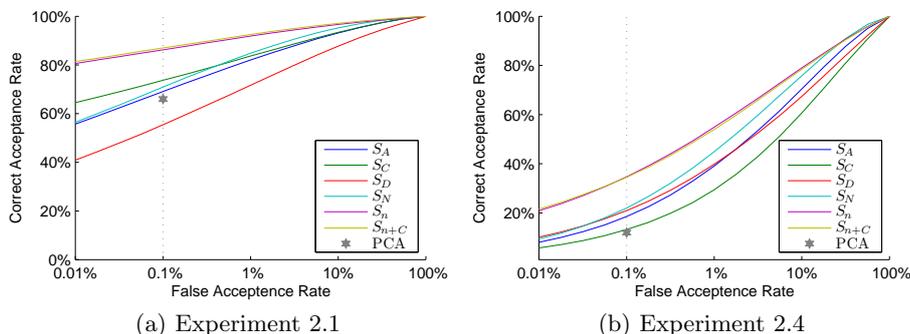


Fig. 3. FRGC ROC III curves: *Dependency of the verification result on the similarity functions for (a) experiment 2.1 and (b) experiment 2.4 of the FRGC database. $\mathcal{G}^{[4]}$ was used throughout.*

The eigenface baseline results 66% for experiment 2.1 and 12% for experiment 2.4 provided by the FRGC database [7] are given in *correct acceptance rates* (CAR) or *verification rates* (VR) at 0.1% *false acceptance rate* (FAR). For comparison with the baseline results, we use the most challenging mask ROC III (cf. [7]), throughout.

The resulting ROC curves are given in Figure 3. Astonishingly, even for experiment 2.1 as given in Figure 3(a) comparing controlled images only, the S_n function already works better than the S_C function, and the combination S_{n+C} outperforms any other function. Generally, $\mathcal{G}^{[4]}$ performed best. The highest verification rate of 87% is achieved by S_{n+C} on $\mathcal{G}^{[4]}$, but this time exceeding S_n on $\mathcal{G}^{[4]}$ only insignificantly. The verification rates for the more demanding problem of comparing uncontrolled with controlled images are presented in Figure 3(b). There, the improvement through the disparity corrected Gabor phase difference S_n is huge compared to any amplitude-only similarity function (VR of 34.7% vs. 18.9% at 0.1% FAR).

5 Discussion

In this paper, we extend the well-known Gabor jet disparity similarity function to allow a higher disparity of the offset positions. Using this estimated disparity, we propose several novel Gabor phase based similarity functions, which turn out to be very useful for recognizing faces photographed with strong or uncontrolled lighting conditions. Especially, images with higher resolutions seem to contain more useful Gabor phase information. Apparently, responses of lower frequency Gabor wavelets hold more illumination specific values, while higher frequency waves code the identity. When the resolution of the original images is sufficient, disparity corrected Gabor phase differences are also beneficial for face recognition under uncontrolled illumination.

Although the proposed system uses a simple average of Gabor jet similarities, it already outperforms state-of-the-art algorithms that use large training sets. Advanced machine learning techniques could further improve face recognition accuracy of our method.

Acknowledgments

We gratefully acknowledge funding from the German Research Foundation (WU 314/2-2 and WU 314/5-2). Portions of the research in this paper use the CAS-PEAL face database collected under the sponsorship of the Chinese National Hi-Tech Program and IS VISION Tech. Co. Ltd. [1].

References

1. Gao, W., Cao, B., Shan, S., Chen, X., Zhou, D., Zhang, X., Zhao, D.: The CAS-PEAL large-scale Chinese face database and baseline evaluations. *IEEE Transactions on Systems, Man, and Cybernetics* 38, 149–161 (January 2008)
2. Günther, M.: Statistical Gabor Graph Based Techniques for the Detection, Recognition, Classification, and Visualization of Human Faces. Ph.D. thesis, Fak. Informatik und Automatisierung, TU Ilmenau, Germany (2011), ISBN: 987-3-8440-0955-2
3. Günther, M., Würtz, R.P.: Face detection and recognition using maximum likelihood classifiers on Gabor graphs. *International Journal of Pattern Recognition and Artificial Intelligence* 23(3), 433–461 (2009)
4. Haufe, D.: Einfluss der Bildauflösung auf Gesichtserkennung durch Graphenvergleich. BSc thesis, Faculty ET-IT, Ruhr-University Bochum (July 2008)
5. Jiménez, D.G., Bicego, M., Tangelder, J., Schouten, B., O. Ambekar, O., Alba-Castro, J., Grosso, E., Tistarelli, M.: Distance measures for Gabor jets-based face authentication: A comparative evaluation. In: *Proc. Intl. Conf. Advances in Biometrics*. pp. 474–483. Springer (2007)
6. Lades, M., Vorbrüggen, J.C., Buhmann, J., Lange, J., von der Malsburg, C., Würtz, R.P., Konen, W.: Distortion invariant object recognition in the dynamic link architecture. *IEEE Transactions on Computers* 42(3), 300–311 (1993)
7. Phillips, P., Flynn, P., Scruggs, T., Bowyer, K., Worek, W.: Preliminary face recognition grand challenge results. In: *Proceedings 7th International Conference on Automatic Face and Gesture Recognition*. pp. 15–24 (2006)
8. Stratmann, B.: Einfluss der Graphenstruktur auf die Leistung eines Gesichtserkennungssystems. BSc thesis, Fak. ET-IT, Ruhr-University Bochum (January 2010)
9. Theimer, W.M., Mallot, H.A.: Phase-based binocular vergence control and depth reconstruction using active vision. *CVGIP: Image Understanding* 60(3), 343–358 (1994)
10. Wiskott, L., Fellous, J.M., Krüger, N., Malsburg, C.: Face recognition by elastic bunch graph matching. *IEEE Trans. PAMI* 19, 775–779 (July 1997)
11. Zhang, W., Shan, S., Gao, W., Chen, X., Zhang, H.: Local Gabor binary pattern histogram sequence (LGBPHS): A novel non-statistical model for face representation and recognition. In: *Proc. ICCV*. vol. 1, pp. 786–791. IEEE (2005)
12. Zhang, W., Shan, S., Qing, L., Chen, X., Gao, W.: Are Gabor phases really useless for face recognition? *Pattern Analysis & Applications* 12, 301–307 (2009)