Finding parameters for Bayesian association of corresponding points with Gabor feature based likelihood

Miika Toivanen a,*, Jouko Lampinen a

*Laboratory of Computational Engineering, Helsinki University of Technology, P.O.Box 9023, 02015 HUT, Finland

Abstract

In feature based object recognition an essential step is finding the locations of the local object features in the image by maximizing a feature similarity function. We study the mapping of a feature similarity function to a likelihood function that reflects the actual probabilities related to associating the object point in a certain location in the image. We adopt probabilistic approach for the feature matching, and suggest a model that contains two adjustable parameters: a steepness parameter of the likelihood function and a threshold parameter related to the probability of the feature appearing in the image. We analyze the effect of the parameters to the feature association, and suggest practical rules for setting the parameters according to the available information, to maximize the number of correct feature associations.

Key words: Bayesian inference, Gabor filters, occlusion, feature likelihood

1 Introduction

In feature based object recognition an essential step is finding the locations of the local object features in the image. Typically this is done by finding locations that maximize a function measuring the similarity of the features found in the image and the stored object feature representation. Alternatively dissimilarity or error measure can be minimized. The similarity function needs to be robust to the variations in the object details occurring in the application,

* Corresponding author.
Email addresses: miika@lce.hut.fi (Miika Toivanen), jlampine@lce.hut.fi (Jouko Lampinen).

Preprint submitted to Elsevier 4 January 2008
to produce a local maximum in the correct location in the analyzed image. However, there is a factor of the similarity measure that is rarely treated explicitly in the recognition systems, that is, the steepness of the similarity measure as function of decreasing similarity.

If we assume probabilistic, or Bayesian, feature matching, the similarity must be measured in terms of the probability of matching the object point to a given image point, and the steepness of the similarity measure (i.e. the likelihood function) must be consistent with the probability of the object producing the features appearing in the image. In this work we study the mapping of Gabor-filter based feature similarity to a likelihood function that reflects the correct probabilities of feature point associations.

The studied issue of the steepness of the feature similarity function is relevant also to the error minimization type approaches, in addition to fully probabilistic approaches. For example, if the object shape is represented as an elastic grid or graph, the joint optimization of the elastic distortion of the graph and the matched feature similarity requires that the steepnesses of the graph stiffness function and the similarity function are adjusted correctly, essentially in the same way as in the probabilistic approach, to maintain the balance of the two optimized terms. Thus the results of the study are applicable also in error minimization approaches.

In this work we study a common feature similarity measure based on the coefficients of Gabor wavelets (Daugman, 1988). Lades et al. (1993) and Wiskott et al. (1997) were among the first to present such similarity measures in their automatic face recognition system, and have influenced multiple face detection and recognition algorithms, some of which use Bayesian methods. For instance, Yan et al. (2004) formulated the face shape localization problem in the Bayesian framework using integrated global and local texture features that base on Gabor wavelets, and Matsui et al. (2004), Lampinen et al. (2001) and Tamminen and Lampinen (2006) use a Bayesian approach with a prior part controlling the deformation of the feature grid.

Object recognition becomes more difficult in the presence of occlusions, since an additional decision to classify the feature point as missing has to be done. As pointed out by Sullivan et al. (2001), images contain statistical information about where the object is and where it is not; therefore in their Bayesian object localization model they employ learned intensities of both the foreground and the background, allowing occlusion to be handled. One way to tackle the occlusion problem is to include the sum over all the hypothesis about the visibilities of the features into the likelihood of the object being present, and compare that to the likelihood of the object being absent. Once the likelihood ratio exceeds a threshold, the object is inferred to be present, and appropriately choosing the threshold, the approach gives good results, even if some
parts are occluded (Weber et al., 2000; Fergus et al., 2003; Fei-Fei et al., 2003; Mikolajczyk et al., 2006). Murphy-Chutorian et al. (2005) use similar approach with a vocabulary of color and Gabor features; at recognition phase, the bins into which the image is Hough transformed cast weighted votes for the presence of the object, using an optimal detection threshold estimated from the training data. Koeser et al. (2006) present a dense optic flow algorithm using Bayesian framework, where occlusion is handled by iteratively searching the corresponding patches between images. Other approaches include computing the similarity between the edges of the model and affine transformed image (Steger, 2002), utilizing Hausdorff distance (Rucklidge, 1997), and using local Gabor binary pattern histogram sequences (Zhang et al., 2005), which work well for occluded objects, as histogram approaches usually do. Occlusion should be taken into account also in stereo matching algorithms (Zitnick and Kanade, 2000).

Tamminen and Lampinen (2006) presented a formal Bayesian Gabor feature based occlusion model, that associates the node points as belonging to the foreground or to the background, that is, as being in occlusion. Here we use a bit different likelihood function, that works better when there is only one training image available. The association of the nodes raises two parameters - a steepness parameter of the likelihood function and a threshold parameter - and in this paper we present optimal values for them. The problems dealing with finding the locations of the nodes are not covered in this paper; the nodes are assumed to already be matched at the correct locations. The paper is organized as follows: Section 2 describes, how the similarities are mapped into the likelihood with the steepness parameter, Section 3 presents a Bayesian way to associate the nodes using the threshold parameter, Section 4 introduces the image databases used in this work, Section 5 presents an analytical solution for the optimal threshold parameter minimizing the association error in a simplified situation, Section 6 considers setting the steepness parameter, and Section 7 concludes.

2 The likelihood model

We approach the problem of matching object in a novel image using Bayesian probabilistic model. There is a sparse grid of node points $x'$ in a reference image, and we construct the posterior probability distribution of their location $x$ in the test image $I$. In the Bayesian formulation the posterior distribution is expressed with the likelihood term and the prior term:

$$p(x|x', I) \sim p(I|x, x')p(x|x'),$$  \hspace{1cm} (1)
where \( p(x|x', I) \) is the posterior probability of \( x \), given the reference grid \( x' \) and the test image \( I \), \( p(I|x, x') \) is the probability of observing \( I \), giving \( x \) and \( x' \), and \( p(x|x') \) is the prior probability for the shape of the node grid, for which we use a Gaussian distribution, whose mean shape is the reference shape, and covariance is diagonal. Equation (1) is proportional up to a normalization term. In this Section we derive the likelihood model used in this work.

Because the likelihood should yield high values for corresponding image points, we use well-known similarity measures to derive the likelihood. Wiskott et. al. (Wiskott, 1995; Wiskott et al., 1997) have proposed two similarity functions, that base on complex Gabor filter responses (Daugman, 1988). The simpler one is just an inner product between the magnitudes \( a_j \) of the Gabor jet coefficients at the two image points \( x \) and \( x' \), normalized to yield illumination invariance (Lades et al., 1993), that is, the cosine of the angle between the two:

\[
S_a(x, x') = \frac{\sum_j a_j(x)a_j(x')}{\sqrt{\sum_j a_j(x)^2 \sum_j a_j(x')^2}}, \quad S_a(x, x') \in [0, 1],
\]  

where the index \( j \) runs over the jet coefficients. The similarity (2), however, ignores the complex phase \( \phi_j \), which carries further information on the accurate location, as pointed out in (Wiskott et al., 1997) and (Lampinen et al., 2001). The phase-sensitive similarity function is defined as

\[
S_p(x, x') = \frac{\sum_j a_j(x)a_j(x') \cos(\phi_j(x) - \phi_j(x'))}{\sqrt{\sum_j a_j(x)^2 \sum_j a_j(x')^2}}, \quad S_p(x, x') \in [-1, 1].
\]  

Similarity (3) is rapidly varying as function of distance, whereas (2) is a smooth function. Using the phase can cause problems, if the nodes are not accurately matched. In (Wiskott, 1995; Wiskott et al., 1997) a displacement term was added into the cosine argument to compensate for the rapid variance of the phase for small displacements. In this work we employed DC-free Gabor filter jets with six different orientations \( (\theta = 0, \pi/6, ..., 5\pi/6) \) and three frequencies \( (f = \pi/4, \pi/8, \pi/16) \).

For either similarity function, the likelihood of node location \( x_i \) is defined as

\[
p(I|x_i, x'_i) = \exp(\beta S(x_i, x'_i)),
\]  

where the parameter \( \beta \) controls the steepness of the likelihood. The likelihood is basically a truncated Gaussian distribution for the normalized filter jets. Assuming the likelihoods of individual nodes to be independent of each other,
the likelihood for the whole grid is the product of the individual likelihoods:
\[ p(I|\mathbf{x}, \mathbf{x}') = \prod_i p(I|x_i, x'_i). \]

The likelihood model of Tamminen and Lampinen (2006) differed slightly from ours; they modeled the amplitude, phase and energy of the Gabor jets with a Gaussian distribution, whose parameters were estimated from training data, and therefore omitted the parameter \( \beta \), but instead had a regularization parameter, whose role is similar to \( \beta \). Modelling the Gabor jets in such a way requires a large training set to estimate the parameter. In this work, we are assumed to have only one reference image, and hence we have to settle with such likelihood as (4). Matsui et al. (2004) in their face recognition system used likelihood similar to ours, with similarity (2) and value \( \beta = 5 \), which they however did not justify in any way.

3 Association probability

The matched nodes must somehow be classified as belonging to the foreground or to the background. In this Section we present a probabilistic tool to associate the nodes into the foreground object.

Let us denote with \( V_i \) the event ‘node \( i \) is present in the image’ and with \( \overline{V_i} \) the opposite event. The Bayesian way to associate the nodes is to form the a posteriori visibility status - association probability - of a node \( i \), by marginalizing over the location \( \mathbf{x} \):

\[
P(V_i|\mathbf{x}', I) = \int p(V_i, \mathbf{x}|\mathbf{x}', I) d\mathbf{x} = \int p(V_i|x_i, x'_i, I)p(\mathbf{x}|\mathbf{x}', I)d\mathbf{x}. \tag{5}
\]

Note that the probability of \( V_i \) is independent of other nodes, if the shape of the occlusion is assumed to be unknown, but a special prior based on neighboring node visibility could be used also.

The calculation of the above integration in closed form is impossible, and because the likelihood is difficult to approximate with simple distributions, sampling methods are needed. The association should then be done to each sample, and take the average over the samples (Tamminen and Lampinen, 2006). Here we concentrate only on the point estimates of the distribution, which are the annotated nodes, and ignore the posterior variance. When \( \beta \) increases, the likelihoods get more peaked by nature, and the posterior reduces to a delta function at the point estimate.

Using Bayesian formula, the probability, that the test image point \( x_i \) is associated with the reference node \( x'_i \), can be written as
\[ P(V_i|x_i, x'_i, I) = \frac{p(I|x_i, x'_i, V_i)P(V_i)}{p(I|x_i, x'_i, V_i)P(V_i) + p(I|x_i, x'_i, \overline{V_i})P(\overline{V_i})}, \]  
(6)

where the prior probabilities for a node being present in an image are assumed to be independent of the location \(x_i\) and the reference node \(x'_i\). The likelihood when the node is detected, \(p(I|x_i, x'_i, V_i)\), is the same as equation (4), and \(p(I|x_i, x'_i, \overline{V_i})\) is the likelihood when there is no correspondence between the reference point and the test point. We try not to construct such a likelihood, but use a constant: \(\kappa_i \equiv p(I|x_i, x'_i, \overline{V_i})\). Tamminen and Lampinen (2006) used the mean of the likelihood as value for \(\kappa_i\). The background parts were modelled with a uniform density also in (Weber et al., 2000; Fergus et al., 2003; Fei-Fei et al., 2003; Mikolajczyk et al., 2006; Koeser et al., 2006), who estimate the value of it from the training data.

The final form for the association probability, showing the parameters \(\beta\) and \(\kappa\) explicitly, is thus

\[ P(V_i|x_i, x'_i, I) = \frac{\exp(\beta S(x_i, x'_i))P(V_i)}{\exp(\beta S(x_i, x'_i))P(V_i) + \kappa_i P(\overline{V_i})}. \]  
(7)

The parameter \(\kappa\) can be considered as a threshold, which separates the nodes into foreground and background nodes. It essentially sets the degree of similarity that we require for the foreground nodes, and therefore depends on task. For instance, if we just want a corner of whatever eye (man’s / woman’s / animal’s etc.) to be associated as the foreground node, the value of \(\kappa\) should be rather low, as compared to a situation, where we want the eye corner only of a specific person to be associated as the foreground node. Similar problem can be found in the shape models. For instance, in (Lades et al., 1993) the stiffness of the grid to be matched is controlled by a parameter, according to how rigid the objects to be matched are; for faces, the parameter should allow larger deformations than for rigid office supplies.

In a pure Bayesian framework, nuisance parameters are integrated out. As we don’t have a posterior distribution for parameters \(\beta\) and \(\kappa\), we have to set some values for them, which is what the rest of the paper deals with.

4 Image databases

In this work we used images with pre-annotated feature points, that represent the point estimates of the posterior distribution (1). Three databases were employed: Bio-Id, DTU, and DigiCam. The first two contain images of human faces, and the last contains images of similar traffic signs, taken with a digital camera. We also included a separate database, denoted by Bio-Id*, which
Fig. 1. Databases used in this paper. Artificial background was added to DTU images. The indices of the background nodes in the Bio-Id and Bio-Id* test images were reversed to reduce the number of coincidentally similar background nodes.

contains only images of the same specific person from the Bio-Id database. Besides the annotated feature points, we added node points to random locations in the background. Representative images of each database with the node points are shown in figure 1. The pixel dimensions and the number of foreground and background nodes in different databases are shown in Table 1.

Table 1

<table>
<thead>
<tr>
<th></th>
<th>width</th>
<th>height</th>
<th>N_FG</th>
<th>N_BG</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bio-Id</td>
<td>384</td>
<td>286</td>
<td>20</td>
<td>10</td>
</tr>
<tr>
<td>DTU</td>
<td>640</td>
<td>480</td>
<td>45</td>
<td>13</td>
</tr>
<tr>
<td>DigiCam</td>
<td>640</td>
<td>480</td>
<td>6</td>
<td>9</td>
</tr>
</tbody>
</table>

From task-dependency of κ, it follows, that its optimal value probably differs for different databases; in the Bio-Id and DTU database, we want to assort different people with varying sex, pose and scale together, whereas in the Bio-Id* images, we want to assort the same person with slightly varying pose and scale together, and in the DigiCam images, we want to assort very similar traffic signs with only slight scale and viewpoint changes together.

5 Setting the threshold parameter by minimizing the association error

In decision theory, an expected utility of making a certain decision is calculated by averaging over the utilities for different outcomes with the decision, weighted by the probabilities for those outcomes. The optimal decision is the one that maximizes the expected utility. In this Section, we consider the ex-
pected utility of the optimal decision, in which all the foreground nodes, and
only them, are associated as foreground nodes, and seek a value for the pa-
rameter $\kappa$, that maximizes the utility.

We assume that the utility is one when making the correct association, and
zero for the opposite case. Furthermore, we take the negation of the expected
utility, and end up with the sum association error:

$$E_s(\beta, \kappa) = - \sum_{i \in \mathcal{FG}} P(V_i | x_i, x'_i, I) - \sum_{i \in \mathcal{BG}} [1 - P(V_i | x_i, x'_i, I)]$$

$$= - \sum_{i \in \mathcal{FG}} \frac{\exp(\beta S(x_i, x'_i)) P(V_i)}{\exp(\beta S(x_i, x'_i)) P(V_i) + \kappa P(V_i)} - \sum_{i \in \mathcal{BG}} \frac{\kappa P(V_i)}{\kappa P(V_i) + \kappa P(V_i)},$$

where $\mathcal{FG}$ denotes the foreground nodes (with correspondence), and $\mathcal{BG}$ the
background nodes (in occlusion). The parameter $\kappa$ is same within the nodes.

The absolute minimum (smallest possible value) of $E_s$ is the total number of
nodes with negative sign.

For arbitrary similarity values the minimum of (8) is unattractive to solve
analytically. However, (8) can be minimized with respect to $\kappa$ with relatively
little effort, if we assume all the similarities and prior association probabilities
of foreground nodes to have the same value (denoted by $S_F$ and $P_F$), and
those of the background nodes to equal as well (with values $S_B$ and $P_B$). The
value of $\kappa$, that minimizes $E_s$, is then

$$\kappa^0 = \frac{2 e^{\beta(S_F + S_B)} \gamma_F \gamma_B (N_B - N_F) + 2 \sqrt{N_F N_B \gamma_F \gamma_B} e^{\beta(S_B + S_F)/2} (e^{\beta S_F} P_F(1 - P_B) - e^{\beta S_B} P_B(1 - P_F))}{2 (N_F e^{\beta S_F} \gamma_F (1 - P_B)^2 - N_B e^{\beta S_B} \gamma_B (1 - P_F)^2)}$$

$$\approx \sqrt{\frac{N_B P_B P_F}{N_F(1 - P_F)(1 - P_B)}} e^{\beta(S_B + S_F)/2},$$

where $\gamma_F \equiv P_F(1 - P_F)$, $\gamma_B \equiv P_B(1 - P_B)$, $N_F$ and $N_B$ denote the number of
foreground and background nodes, and the approximation holds when $\beta \gtrsim 1$
and $S_F > S_B$. These conditions are always met in practice.

From now on, we take all the prior association probabilities to be half, in
which case the above estimate reduces to

$$\kappa^0 = \sqrt{\frac{N_B}{N_F}} e^{\beta(S_B + S_F)/2}.$$  \hspace{1cm} (9)

The negation of the probability of associating all the nodes correctly is
Table 2
Means ± standard deviations of the background and foreground similarities in different databases. Subscript \( a \) refers to the phase-insensitive similarity and subscript \( p \) to the phase-sensitive similarity. Due to memory limitations, only every 10th pixel was used.

<table>
<thead>
<tr>
<th></th>
<th>( S_a(BG) )</th>
<th>( S_a(FG) )</th>
<th>( S_p(BG) )</th>
<th>( S_p(FG) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bio-Id</td>
<td>0.65 ± 0.19</td>
<td>0.88 ± 0.065</td>
<td>0.0021 ± 0.29</td>
<td>0.51 ± 0.29</td>
</tr>
<tr>
<td>DTU</td>
<td>0.65 ± 0.16</td>
<td>0.88 ± 0.077</td>
<td>-6.8E-4 ± 0.24</td>
<td>0.53 ± 0.27</td>
</tr>
<tr>
<td>Bio-Id*</td>
<td>0.65 ± 0.19</td>
<td>0.93 ± 0.072</td>
<td>-0.0054 ± 0.29</td>
<td>0.84 ± 0.14</td>
</tr>
<tr>
<td>DigiCam</td>
<td>0.63 ± 0.17</td>
<td>0.96 ± 0.036</td>
<td>5.8E-4 ± 0.24</td>
<td>0.82 ± 0.15</td>
</tr>
<tr>
<td>Rand pics</td>
<td>0.65 ± 0.16</td>
<td>-</td>
<td>0.0010 ± 0.24</td>
<td>-</td>
</tr>
<tr>
<td>All</td>
<td>0.65 ± 0.17</td>
<td>0.90 ± 0.076</td>
<td>-4.1E-4 ± 0.25</td>
<td>0.61 ± 0.28</td>
</tr>
</tbody>
</table>

\[
E_p(\beta, \kappa) = -\prod_{i \in F_B} P(V_i | x_i, x_i', I) \prod_{i \in B} [1 - P(V_i | x_i, x_i', I)]
= -\prod_{i \in F_B} \frac{\exp(\beta S(x_i, x_i'))}{\exp(\beta S(x_i, x_i')) + \kappa} \prod_{i \in B} \frac{\kappa}{\exp(\beta S(x_i, x_i')) + \kappa} ,
\]

whose minimization under the same assumptions as above leads also to estimate (9).

![Histograms of the background similarities in different databases. Top row corresponds to the phase-insensitive and bottom row to the phase-sensitive similarities. On the right are the histograms of all the 60 image pairs.](image)

The logarithm of RHS of equation (9) is linear in \( \beta \):

\[
\log(\kappa^0) = \frac{1}{2}(S_B + S_F)\beta + \frac{1}{2} \log \left( \frac{N_B}{N_F} \right) \equiv w_2\beta + w_1 .
\]

The result is easily interpretable. When there are as many foreground nodes as background nodes, \( \log(\kappa^0) \) is just \( \beta \) times the mean of the background and fore-
Fig. 3. The errors based on summation ($E_s$) and product ($E_p$) of individual node association probabilities, using phase-insensitive similarities. Middle panel shows actually the negation of the logarithm of the negative error surface. The lines show the logarithms of $\kappa(\beta)$ that minimizes the errors, which are also plotted in the right panel, where superscripts $s$ and $p$ correspond to $E_s$ and $E_p$, and 0 to the equation (11). See text for details.

ground similarities (parameter $w_2$), and with unequal number of background and foreground nodes, a small bias (parameter $w_1$) is added.

The formula (11) is valid only when all the foreground similarities are equal, and respectively for the background similarities. In a realistic situation, there is naturally some variation in the similarities. To our knowledge, no exhaustive study of how the Gabor similarities are distributed in natural images has been done; Okada and Lyons (2002), who used Gabor filters in classificating and registrating stained rat brain slices, studied the similarity distributions in their images. We computed the background similarities of 60 image pairs.

Out of each database we chose 10 image pairs, and in addition we picked from the World Wide Web 60 random varied images, containing for instance artistic drawings and text. The images formed 30 pairs, so that 10 nodes were randomly chosen from the other images, and the similarities with the other images were computed. The histograms of the phase-insensitive and phase-sensitive background similarities in different databases are depicted in figure 2, and the statistics are tabulated in table 2, together with the ones of the foreground similarities at the corresponding annotated feature points (excluding, of course, the WWW images). The background similarities are approximately equal between the databases, and the phase-sensitive similarities are symmetric about zero, as expected due to the cosine term. The difference in the foreground similarities naturally reveals the difference in the in-class variation between the databases, which is low in the Bio-Id* and DigiCam images. The phase-insensitive background histograms resemble that in (Okada and Lyons, 2002).

To assess the effect of variation in similarities on the parameters $w_2$ and $w_1$, we computed the error surfaces (8) and (10), so that the background and foreground similarities were randomly drawn from the distributions in figure 2. As an illustration, $E_s(\beta, \kappa)$ and $E_p(\beta, \kappa)$ with $N_F = 60$ foreground and $N_B = 40$ background phase-insensitive similarities are represented in figure
3. The absolute minimum of $E_s$ is $N_F + N_B = 100$. On the right panel, the theoretical value $\log(\kappa^0)$ is equation (11), into which the mean values from the bottom row of table 2 was put. At least for these particular $N_F$ and $N_B$ values, and for these similarity values drawn, the curves seem to be in good agreement with each other.

![Graphs showing quartiles of linear fit parameters $w_2$ and $w_1$ as a function of ratio of number of background to foreground nodes.](image)

Fig. 4. Quartiles of the linear fit parameters $w_2$ and $w_1$, as a function of ratio of number of background to foreground nodes. Superscripts $s$ and $p$ correspond to minimizing sum error $E_s$ and product error $E_p$, and superscript 0 denote the formula (11). Top row corresponds to the phase-insensitive and bottom row to the phase-sensitive similarities.

A first order polynomial was fitted to the $\log \kappa(\beta)$ curves in the range $\beta = [10, 20]$. This was firstable done 200 times, so that the similarities were randomly drawn each time, and this was then repeated for different fractions of the number of background and foreground nodes, $N_B/N_F$. The parameters of linear fits - slope $w_2$ and intercept $w_1$ - are plotted against a reasonable range of $N_B/N_F$ in figure 4, together with the theoretical parameters of formula (11), using again the values of the bottom row of table 2. With large $N_B/N_F$ values, $E_s$ with phase-insensitive similarities tend to lack local minimum (at least in the range we used) and $\kappa_{\min}(\beta)$ is just a vertical line on the right edge, which results in $w_2$ being zero and $w_1$ being $\max[\log(\kappa)]$ in those cases. The theoretical parameters of formula (11) are in close agreement with the computed ones, especially in the vicinity of $N_B/N_F = 1$. What is remarkable, is that with the phase-insensitive similarities, the $w_2$ parameter is approximately constant w.r.t. $N_B/N_F$, while with the phase-sensitive similarities, this is not so (compare the changes in the $w_2$-axis in the two cases).
The association error surfaces in $(\kappa, \beta)$ space were computed for several annotated image pairs, chosen from the databases presented in section 4. The similarities between the pre-annotated node points, log $(\kappa_{\text{min}})$ curves for $E_s$ and $E_p$, and the parameters $w_2$ and $w_1$ were computed. These are tabulated in table 3, together with the theoretical parameters of formula (11), using the values of table 2 for the similarities.

In order to be able to use the equation (11), one has to insert some values for $S_B$, $S_F$, and the ratio $N_B/N_F$, but these are usually unknown a priori, although it might be possible to learn them from the data. For $S_B$ and $S_F$, the mean values from table 2 could be inserted, and assuming $N_B = N_F$ results in simplified formulas:

\[ \kappa^0 \approx \exp(\beta) \] (12a)

\[ \kappa^0 \approx \exp(0.305 \beta) \] (12b)

### 6 Setting the steepness parameter

Fig. 5. The effect of altering $\beta$ parameter. Foreground nodes are marked with blue plus signs, background nodes with red crosses. The value of $\kappa$ is kept constant at $\kappa = \exp(5)$. Besides the current value of $\beta$, also summation and product errors are shown above each figure.

Having formulated the procedure for computing $\kappa$, the errors $E_s$ and $E_p$ turn into 1-dimensional, being a function of $\beta$ only. The way in which the errors depend on $\beta$ is being illustrated in figure 5 using an artificial example, where one of the background similarities is larger than two of the foreground similarities, and $\kappa$ is constant. The attempt to keep as many of the foreground association probabilities as close to one as possible, while keeping as many of the
Table 3

The parameters of linear fit for different databases. $w_s^2$ and $w_s^1$ denote the slope and intercept of the first order polynomial for log($\kappa_{\text{min}}$) minimizing the sum error $E_s$, and $w_p^2$ and $w_p^1$ are those for the product error $E_p$, respectively. The numbers above the horizontal line were received using phase-insensitive similarities, and vice versa. In parentheses are the number of image pairs.

<table>
<thead>
<tr>
<th></th>
<th>Bio-Id(46)</th>
<th>DTU(18)</th>
<th>Bio-Id*(20)</th>
<th>DigiCam(17)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$w_s^2$: mean ± std</td>
<td>0.75 ± 0.10</td>
<td>0.79 ± 0.06</td>
<td>0.84 ± 0.08</td>
<td>0.85 ± 0.04</td>
</tr>
<tr>
<td>$w_s^1$: mean ± std</td>
<td>0.78 ± 0.04</td>
<td>0.79 ± 0.03</td>
<td>0.85 ± 0.06</td>
<td>0.87 ± 0.04</td>
</tr>
<tr>
<td>$\frac{1}{2}(S_B + S_F)$</td>
<td>0.77</td>
<td>0.77</td>
<td>0.79</td>
<td>0.80</td>
</tr>
<tr>
<td>$w_p^2$: mean ± std</td>
<td>-1.78 ± 1.46</td>
<td>-2.37 ± 0.70</td>
<td>-1.36 ± 0.47</td>
<td>0.14 ± 0.23</td>
</tr>
<tr>
<td>$w_p^1$: mean ± std</td>
<td>-0.77 ± 0.18</td>
<td>-1.05 ± 0.15</td>
<td>-0.78 ± 0.24</td>
<td>-0.05 ± 0.22</td>
</tr>
<tr>
<td>$\frac{1}{2} \log(N_B/N_F)$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The errors were again calculated for each image pair in order to investigate, if $E(\beta)$ has a local minimum. Also the mean of the likelihood over each test image, advocated as value for $\kappa$ in (Tamminen and Lampinen, 2006), was computed and compared against the proposed $\kappa$. Investigating the magnitude of $E(\beta)$ also serves as a check, how ’dangerous’ it is to insert a ’wrong’ value for $\kappa$.

The average summation errors are plotted in figures 6 and 7. The product errors behave in a similar way, and are not shown here due to the shortage of space. With small $\beta$, all the parameters $\kappa$ produce approximately equal errors. With the Bio-Id* and DigiCam databases, similarity exists also with large $\beta$ values, whereas with the Bio-Id and DTU databases, the proposed method is superior. Also there is usually little difference between using equation (11) with the ’exact’ values from table 2, and the approximations (12), and in some cases the errors are very close to the smallest possible error $E_{\text{min}}$. When $\beta \gtrsim 20$, the errors typically decrease only slowly. It should also be noted,
that with the Bio-Id* and DigiCam images, using phase-sensitive similarities results in slightly smaller association errors.

For certain search methods or sampling methods, the location of the grid is easier to find with low-gradient likelihoods. If the aim is also to associate the nodes, then a method that can deal with large $\beta$ should be used.

![Graphs showing one-dimensional association errors for different databases with phase-insensitive similarities.](image)

**Fig. 6.** The one-dimensional association errors for different databases with phase-insensitive similarities. $E^0$ stands for equation (11); $E^A$ stands for the estimates (12); $L_M$ denotes the mean likelihood; $E_{\text{min}}$ corresponds to using $\kappa_{\text{min}}$. For sake of clarity, only the means over the image pairs are shown.

### 7 Conclusions

In this paper we have mapped the similarity of the corresponding points into the likelihood, using parameter $\beta$. The similarities are based either only on the absolute values of the complex Gabor jets, or also on the phases. The association probability introduced the parameter $\kappa$, for which an optimal value was analytically derived in an oversimplified situation, where all the foreground similarities and all the background similarities had same values. The parameters of the linear fit for $\kappa$ minimizing the association error were close to the parameters of the analytical formula, at least the slope parameter, which dominates with large values of $\beta$. Finally we considered the effect of chang-
Fig. 7. Same results as in figure 6, but with phase-sensitive similarities. 

ing $\beta$, and expressed the errors as a function of $\beta$. The results show that 
the proposed $\kappa$ works equally well or better than the previously used mean 
likelihood value, depending on the $\beta$, and often differs little from the lowest 
possible error. With the proposed $\kappa$ there seems to be no local minimum for 
$\beta$ - the errors decrease as $\beta$ increases. The results furthermore show, that the 
phase-sensitive similarity works on average a little bit better than the phase-
insensitive similarity, although the results of table 3 somewhat advocated the 
phase-insensitive similarity.

Based on these results, when using likelihood (4) for matching and associating 
the nodes with equation (7), we recommend using as parameter $\kappa$ the formula 
(11), when the ratio of background and foreground nodes and the level of the 
foreground similarity is known, or might be learned from the data. When this 
is not the case, the recommend using the approximation (12).

References

Daugman, J., 1988. Complete discrete 2-D Gabor transforms by neural net-
works for imageanalysis and compression. Acoustics, Speech, and Signal 
Processing [see also IEEE Transactions on Signal Processing], IEEE Trans-
actions on 36 (7), 1169–1179.


