Invariance Properties of Gabor Filter Based Features – Overview and Applications

Joni-Kristian Kamarainen*, Ville Kyrki, *Member, IEEE*, Heikki Kälviäinen, *Member, IEEE*

Abstract— **For almost three decades the use of features based on Gabor filters has been promoted for their useful properties in image processing. The most important properties are related to invariance to illumination, rotation, scale and translation. These properties are based on the fact that they are all parameters of Gabor filters themselves. This is especially useful in feature extraction, where Gabor filters have succeeded in many applications, from texture analysis to iris and face recognition. This study provides an overview of Gabor filters in image processing, a short literature survey of the most significant results, and establishes invariance properties and restrictions to the use of Gabor filters in feature extraction. Results are demonstrated by application examples.**

Index Terms— **Gabor filters, feature extraction, invariance, object detection, object recognition, image processing.**

I. INTRODUCTION

INVARIANCE of features has been an active area in object recognition research for decades. By invariance not only are features meant which are invariant to a set of geometric transforms but also methods to perform object detection regardless of pose and imaging conditions using features which are not invariant. Invariance requirements are application dependent. In the recognition of 2-D rigid objects, for example, illumination, orientation, scale, and translation invariances are the most typical requirements. General approaches, such as moment invariants, do exist, but have problems in practice since they require precise segmentation and uniform lighting. Gabor features have instead been utilized since they can preserve pose information in the feature space. This study presents an overview and a literature survey of Gabor filters in feature extraction, their invariance properties, and describes results in applications utilizing Gabor features.

Works utilizing Gabor filters often refer to the original contribution by Dennis Gabor in 1946 when he proposed a represention of signals as a combination of elementary functions [1]. Gabor's work was a continuation and partly parallel to the works of Nyquist [2] and Shannon [3] founding the theory of communication. Since then Gabor functions have been deployed in many areas of signal processing and to list some of the most influential works which contributed to this research, credit must be given to the detailed analysis of Gabor's expansion by Bastiaans [4], [5], [6], the introduction of the two dimensional (2-D) counterpart of a Gabor elementary function by Granlund [7], the generalization to two dimensions and the 2-D function as a model of a simple cell in the mammalian visual cortex by Daugman [8], and the wavelet treatment of Gabor filters by Lee [9].

In the overview part of this study the fundamental results involved in the discovery and development of Gabor filters in signal processing are described. In addition, the study formulates invariance properties in more interpretable forms and their use is demonstrated in applications. The Gabor research community lacks an adequate overview of the field, and for this reason, the authors have collated an extensive list of references. Gabor researchers have combined results from many different contexts and consequently the terminology used may be confusing to researchers who are not aware of the context they should work with; Gabor filter, Gabor expansion, Gabor transform, Gabor jet, Gabor frame, or Gabor wavelet? Despite the fact that this study begins from the filters in one dimension, the study is mainly relevant to image processing. The image processing community seems to be more familiar with Gabor filters compared to other research areas, which is probably due to the physiological fact that, as Daugman pointed out, the 2-D Gabor filters are good models of the simple cells in the mammalian visual system [8]. This qualitative fact has promoted the use of Gabor filters for twenty years and is still mentioned in almost every article on the subject. Nevertheless, the physiological issues are only briefly considered in this study which attempts to provide motivations for Gabor filter use with a more quantitative rationale.

II. SURVEY AND OVERVIEW

A. Signals in two domains

In signal analysis, and especially in image processing, feature extraction plays a central role in "Knowing what is where?". If one can be convinced that frequency content provides the information about "what", and spatial coordinates in 2-D provide the information about "where", a connection to the uncertainty principle and Gabor's work can be established. First what makes these two domains, time and frequency, special should be considered and is it reasonable to restrict the study to these domains alone.

1) Time: With one-dimensional (1-D) signals it is natural to analyze their variation in time since most signals are measured in time domain and time itself is fundamental. However, it should be noted that for some other signals, e.g., two dimensional (2-D) images, the characteristics of the dimensions are not necessarily comparable to time.

All authors are with the Laboratory of Information Processing, Department of Information Technology, Lappeenranta University of Technology, P.O. Box 20, FIN-53851 Lappeenranta, Finland. E-mail: Joni.Kamarainen@lut.fi, Ville.Kyrki@lut.fi, Heikki.Kalviainen@lut.fi .
* Corresponding author
EDICS: 2-ANAL, 2-OTHB

2) Frequency: Waveforms, optical, electrical, or acoustical, and their Fourier spectra are appreciated equally as physically picturable and measurable entities. Gabor was also concerned about reasons for the use of frequency domain other than the elementary mathematical properties and strong theory of the harmonic wave functions. In his study Gabor pointed out that some other selection of orthogonal basis, like Bessel functions, would not provide spectral components whose number is proportional to the length of a specific time interval [1].

3) Uncertainty: To analyze signals, an operator is needed that performs the analysis simultaneously in both time and frequency domains and provides information of localized time and frequency events. Let $\psi(t)$ denote a function in terms of time and $\Psi(u)$ its Fourier transform in terms of frequency. These two functions are connected via the Fourier transform pairs

$$\Psi(f) = \mathfrak{F}\{\psi(t)\}, \ \ \psi(t) = \mathfrak{F}^{-1}\{\Psi(f)\} \ . \tag{1}$$

Clearly function ψ is situated in both domains, but since it is well known that any compactly supported function cannot have a finite Fourier transform and vice versa [10], there is always uncertainty in the time and frequency locations of ψ.

Bandwidth has been widely studied in communication systems and it seems that even in this field no single definition suffices [11], [12]. The bandwidth definition that Gabor found useful was the root mean square (r.m.s.) bandwidth. A similar r.m.s. measure can be applied to uncertainty in time, i.e., the r.m.s. duration. In literature these two measures have been referred to as Gabor bandwidths [11], [13]. The uncertainty in time, the time duration Δt, can be defined as ([8])

$$\Delta t = \sqrt{\frac{\int_{-\infty}^{\infty}(t-\mu_t)^2\psi(t)\psi^*(t)dt}{\int_{-\infty}^{\infty}\psi(t)\psi^*(t)dt}}, \ \ \mu_t = \frac{\int_{-\infty}^{\infty}t\psi(t)\psi^*(t)dt}{\int_{-\infty}^{\infty}\psi(t)\psi^*(t)dt} \tag{2}$$

where μ_t is the centroid of ψ in time and ψ^* denotes the complex conjugate. A similar uncertainty Δf for frequency can be defined, and then, these two uncertainties are connected via the uncertainty principle as

$$\Delta t \Delta f \geq \frac{1}{4\pi} \ . \tag{3}$$

There are several formulas for the Fourier transform varying by constant factors, and thus, the uncertainty value is sometimes replaced by $1/2$, but these all have an identical interpretation [10], [8], [1].

It is possible to derive a function ψ for which the product $\Delta t \Delta f$ assumes the smallest possible value, i.e., for which the inequality in (3) turns into an equality: *"The signal which occupies the minimum area $\Delta t \Delta f = \frac{1}{4\pi}$ is the modulation product of a harmonic oscillation of any frequency with pulse of the form of a probability function"* [1]

$$\psi(t) = e^{-\alpha^2(t-t_0)^2}e^{j2\pi f_0 t + \phi} \tag{4}$$

where α is the sharpness (time duration / bandwidth) of a Gaussian, t_0 the centroid of the Gaussian, f_0 is the frequency of the harmonic oscillations, and ϕ denotes the phase shift of the oscillation. The Gabor elementary function in (4) has the Fourier spectrum

$$\Psi(f) = \sqrt{\frac{\pi}{\alpha^2}}e^{-(\frac{\pi}{\alpha})^2(f-f_0)^2}e^{-j2\pi t_0(f-f_0)+\phi} \ . \tag{5}$$

It is straightforward to show that $\Delta t \Delta f = \frac{1}{4\pi}$. The form of the function, the Gabor elementary function ψ, that minimizes the uncertainty and turns the inequality (3) into equality is now defined and the use and properties of the function will be considered next.

B. Gabor expansion

Signal expansion, synthesis, was Gabor's main motivation when he proposed the use of Gabor elementary functions in (4) as expansion functions which would optimally represent the time-frequency content. It should be noted that the expansion functions do not have to constitute an orthogonal basis as typically assumed in wavelet or Fourier transforms, but an unconditional basis, a frame, may succeed as well [14]. In the Gabor expansion, a signal $s(t)$ is represented as a sum of Gabor elementary functions $\psi_{kl}(t)$ multiplied by expansion coefficients a_{kl} as

$$s(t) = \sum_{k,l} a_{kl}\psi_{kl}(t), \ \ \psi_{kl}(t) = e^{-\alpha^2(t-kt_0)^2}e^{j2\pi lf_0 t + \phi} \tag{6}$$

where k denotes time shifts and l frequency shifts of the Gabor function. The expansion coefficients a_{kl} must be computed using a biorthogonal function set [4], [14]. The biorthogonality is trivial for orthogonal bases but more complicated for frames. It has been shown that Gabor's method to construct the biorthogonal set is impractical [?] and more stable methods have been introduced by Bastiaans [4], [5], [6]. A unique solution can be found only for a special case of the parameters k and l, namely , when the signal is critically sampled ($t_0 f_0 = 2\pi$), but even then the convergence may be unstable [4], [5], [15]. For the values $t_0 f_0 < 2\pi$ the Gabor frame is overcomplete and the biorthogonal set is no longer unique, nevertheless, various methods for constructing the set of biorthogonal functions have been proposed [6], [16], [17].

Despite the ambiguities the Gabor expansion has advantages compared to other time-frequency methods [18] and it has been successfully applied to analysis and used in applications [19], [20]. Lately, inspired by wavelets the Gabor expansion has been analyzed in the context of filter banks [21], [22] and the Gabor theory (e.g. [23]) has been extended towards wavelets (e.g. framelets in [24]).

C. 1-D Gabor filter

Next, instead of signal synthesis as in the expansion, the Gabor elementary function is utilized as a signal analyzing filter. To be consistent with linear filter theory a few justifications need to be made to (4) in order to define a proper form of the Gabor filter function [25]. First, an origin centered form ($t_0 = 0$) is preferred for the convolution. There is no evidence that any specific phase (ϕ) would be more beneficial than any other, so the phase shift can be removed to define the Gabor filter function in a more compact form

$$\psi(t) = e^{-\alpha^2 t^2}e^{j2\pi f_0 t} \ . \tag{7}$$

Now, response of the linear filter in (7) at some location t_1 for some signal $\xi(t)$ can be calculated with the convolution

$$resp(t_1) = \psi(t_1) * \xi(t_1) = \int_{-\infty}^{\infty} \psi(t_1 - t)\xi(t)dt \quad (8)$$

which is similar to the continuous short-time Fourier transform with a Gaussian window function [26]. The only difference is the time dependent phase factor $e^{j2\pi f_0 t_1}$ which is caused by the assumption of the $0°$ phase angle in the filter centroid.

1) Multi-resolution analysis filter: The filter in (7) has the same effective widths, duration and bandwidth, regardless of the central frequency f_0. However, a multi-resolution analysis (MRA) compatible form can be introduced to make the filters on different frequencies behave as scaled versions of each other [14]. Intuitively this makes sense since it is reasonable to inspect the same events but in different scales; the higher the frequency the finer the details. The time duration of the filter should depend on the central frequency f_0 [27] which can be accomplished by setting the frequency normalized bandwidth parameter

$$\gamma = \frac{|f_0|}{\alpha} \quad . \quad (9)$$

2) Response normalization: From the Fourier transformed Gabor function in (5) it can be seen that the maximum response for a complex signal is $\sqrt{\pi/\alpha^2}$, and thus, its inversion can be used as a normalization term. Finally, a normalized 1-D Gabor filter function can be defined ($f_0 = f$) as

$$\psi(t) = \frac{|f|}{\gamma\sqrt{\pi}} e^{-(\frac{f}{\gamma})^2 t^2} e^{j2\pi f t} \quad (10)$$

which is referred to as the 1-D Gabor filter hereafter. The filter has a Fourier transform

$$\Psi(u) = e^{-(\frac{\gamma\pi}{f})^2(u-f)^2} \quad (11)$$

where u denotes frequency.

D. 2-D Gabor filter

It is straightforward to generalize the filter in (10) to two dimensions where the time variable t is replaced by the spatial coordinates (x, y) and the frequency variable u by the frequency variable pair (u, v) and this was presented in the late 70's. If similar studies in optics are not considered (e.g. [28], [29]), the major effort to the development and use of 2-D Gabor filters has been made in image processing and especially in feature extraction. In 1978 Granlund proposed the form of a general picture processing operator which corresponds to a 2-D Gabor filter and was derived directly from requirements of image processing [7]. It is noteworthy that Granlund addressed many properties, such as the octave spacing of the frequencies, that were reinvented later for Gabor filters.

Daugman later showed a surprising equivalence between a structure based on the 2-D Gabor functions and the organization and the characteristics of the mammalian visual system [8], [30]. He also defined similar uncertainty measures as in (2) for the 2-D case in terms of Δx, Δy, Δu, and Δv for which it holds that

$$\Delta x \Delta u \Delta y \Delta v \geq \frac{1}{16\pi^2} \quad . \quad (12)$$

The most general form of the functions that achieve the lower bound of the uncertainty inequality in (12) is

$$e^{-(Ax^2 + Bxy + Cy^2 + Dx + Ey + F)} \quad (13)$$

where $B^2 < 4AC$ and D, E, and F are complex [8], but for practical use certain simplifications are typically made. The simplifications are particularly motivated if the construction of the 2-D Gabor filters is done according to biological evidence (e.g. [8]), but they can also be justified based on linear filter theory [25].

Furthermore, according to the MRA form in (10) substitutions $\alpha = |f_0|/\gamma$ and $\beta = |f_0|/\eta$ can be made and the corresponding 2-D Gabor filter can be defined as

$$\psi(x, y) = \frac{f^2}{\pi\gamma\eta} e^{-(\frac{f^2}{\gamma^2}x'^2 + \frac{f^2}{\eta^2}y'^2)} e^{j2\pi f x'}$$
$$x' = x\cos\theta + y\sin\theta, \ y' = -x\sin\theta + y\cos\theta \ . \quad (14)$$

where f is the central frequency of the filter, θ the rotation angle of the Gaussian major axis and the plane wave, γ the sharpness along the major axis, and η the sharpness along the minor axis (perpendicular to the wave). Note that the center of the filter is defined in polar coordinates with parameters (f, θ). In the given form, the aspect ratio of the Gaussian is $\lambda = \eta/\gamma$. The normalized 2-D Gabor filter function in the frequency domain is

$$\Psi(u, v) = e^{-\frac{\pi^2}{f^2}(\gamma^2(u'-f)^2 + \eta^2 v'^2)}$$
$$u' = u\cos\theta + v\sin\theta, \ v' = -u\sin\theta + v\cos\theta \ . \quad (15)$$

The effects of the parameters, interpretable via the Fourier similarity theorem, are demonstrated in Figs. 1 and 2.
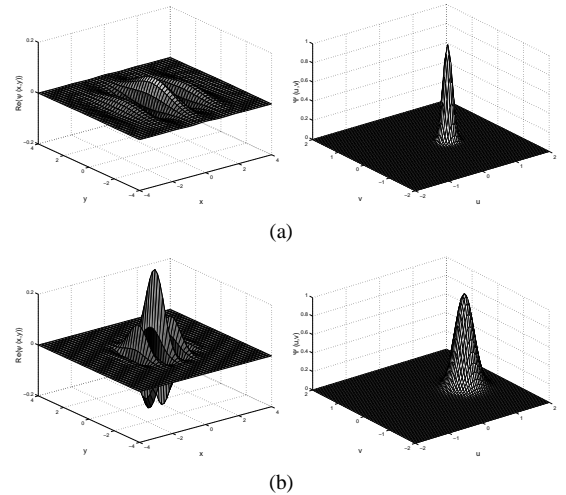

(a)

(b)

Fig. 1. Examples of 2-D Gabor filters in spatial and frequency domains: (a) $f = 0.5$, $\theta = 0°$, $\gamma = 1.0$, $\eta = 1.0$; (b) $f = 1.0$, $\theta = 0°$, $\gamma = 1.0$, $\eta = 1.0$.

III. DISCRETE FILTERS

To transform continuous domain entities to the discrete domain one always needs to be sure that they are represented with sufficient accuracy in order to reapply the continuous domain results. The discrete domain has been addressed in
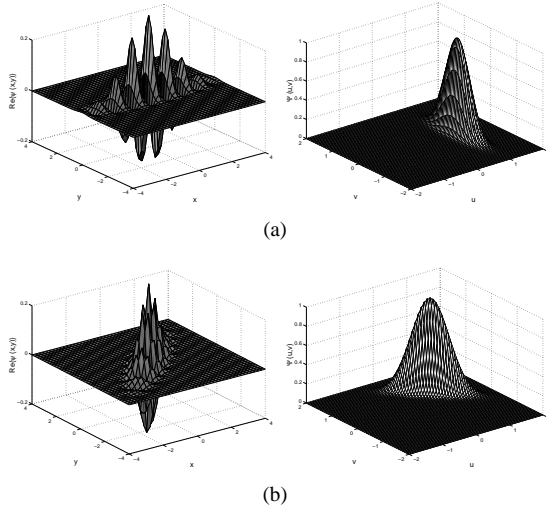
Fig. 2. Examples of 2-D Gabor filters in spatial and frequency domains: (a) $f = 1.0$, $\theta = 0°$, $\gamma = 2.0$, $\eta = 0.5$; (b) $f = 1.0$, $\theta = 45°$, $\gamma = 2.0$, $\eta = 0.5$.

several studies (e.g. [5], [6], [16], [15]), but the main concern has been the construction of biorthogonals, not the sampling and quantization of Gabor and biorthogonal functions or the expansion coefficients.

Real signals are usually strictly amplitude limited and quantization is not an issue. In applications, a proper construction depends more on the sampling theorem, which has been only briefly discussed for Gabor filters [31], [32], [33]. By obeying the sampling theorem an accurate and aliasing free construction of Gabor filters can be made in the frequency domain. However, the avoidance of aliasing only guarantees a proper sampling of the filter, but it does not say anything about a proper filter in the time domain. In the time domain it would be beneficial if the filter envelope falls completely in discrete bins of a filter representation, that is, the filter values are negligible outside a limited length discrete filter [25].

As any continuous function, the Gabor filter in (10) can be sampled as a discrete sequence

$$\psi(n) = \psi(Tt) \ . \tag{16}$$

Now, a discrete 1-D normalized Gabor filter can be defined as

$$\psi(n) = \frac{|f|}{\gamma\sqrt{\pi}} e^{-(\frac{f}{\gamma})^2 n^2} e^{j2\pi f n} \tag{17}$$

where f denotes discrete frequencies between $\frac{n_{min}}{L}$ and $\frac{n_{max}}{L}$, where n_{min} and n_{max} correspond to the bins of the first and last discrete frequencies and L is the filter length.

To ensure that a filter is properly inside the frequency range, that is, at most the proportion $1 - p_f$ is aliased, it must hold that (positive frequencies)

$$\sum_{n=1}^{n_{max}-1} \Psi(n) \geq p_f \sum_{n=-\infty}^{\infty} \Psi(n), \ 0 < f < 0.5 \tag{18}$$

which guarantees that the following condition for the continuous case holds: [25]

$$\int_0^{\frac{n_{max}}{L}} \Psi(u)du \geq p_f \int_{-\infty}^{\infty} \Psi(u)du \ . \tag{19}$$

A proper filter construction in the frequency domain can be guaranteed by satisfying [25]

$$\frac{1}{2}\operatorname{erf}(\gamma\pi) + \frac{1}{2}\operatorname{erf}\left(\frac{\gamma\pi}{|f|}(\frac{1}{2} - \frac{1}{L}) - \gamma\pi\right) \geq p_f \tag{20}$$

where erf is the error function

$$\operatorname{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt \ . \tag{21}$$

Eq. (20) provides a sufficient sampling of the filter in the time domain, but the filter envelope may spread beyond the discrete size of the filter. The effective filter envelope is specified by the modulating Gaussian, and thus, a restriction

$$\sum_{n_{min}}^{n_{max}} e^{-(\frac{f}{\gamma})^2 n^2} \geq p_t \sum_{-\infty}^{\infty} e^{-(\frac{f}{\gamma})^2 n^2} \tag{22}$$

must be satisfied where the factor p_t corresponds to p_f in the frequency domain. By lower bound estimations a general constraint that holds for odd and even L can be derived and the time domain constraint can be written as

$$\operatorname{erf}\left(\frac{|f|}{\gamma}(\frac{L}{2} - 1)\right) \geq p_t \ . \tag{23}$$

Eq. (23) can also be used to estimate the minimum size L of the filter.

The results can be generalized to two dimensions [25]. To define a strict lower boundary, L can be set equal to the size of the smallest dimension of the 2-D filter, $L = \min\{L_x, L_y\}$, and the filter can be considered in the standard pose where the major axis is along the x-axis ($\theta = 0$). Now, the frequency constraint for the 2-D filter is

$$\frac{1}{4}\left[\operatorname{erf}(\gamma\pi) + \operatorname{erf}\left(\frac{\gamma\pi}{|f|}(\frac{1}{2} - \frac{1}{L}) - \gamma\pi\right)\right] \\ \left[\operatorname{erf}(\eta\pi) + \operatorname{erf}\left(\frac{\eta\pi}{|f|}(\frac{1}{2} - \frac{1}{L}) - \eta\pi\right)\right] \geq p_f \tag{24}$$

and the spatial constraint

$$\operatorname{erf}\left(\frac{|f|}{\gamma}(\frac{L}{2} - 1)\right)\operatorname{erf}\left(\frac{|f|}{\eta}(\frac{L}{2} - 1)\right) \geq p_t \ . \tag{25}$$

Using the discrete form of the Gabor filter in (17) or a corresponding 2-D form, and by satisfying the given constraints in (20) and (23) or in (24) and (25) a proper construction of the Gabor filters can be achieved and reliable results in applications can be expected.

IV. INVARIANCE PROPERTIES

A. Time-frequency features of 1-D signals

The main use of Gabor filters is with 2-D images, but here they are first demonstrated in one dimension for simplicity. Gabor features from a signal $\xi(t)$ can be extracted via the convolution

$$r_\xi(t; f) = \psi(t; f) * \xi(t) = \int_{-\infty}^{\infty} \psi(t - t_\tau; f)\xi(t_\tau)dt_\tau \tag{26}$$

where the frequency f is included as a feature variable. Via the convolution, the Gabor filter works as any linear filter and the response $r_\xi(t; f)$ in (26) can be calculated at any time

instant t and for various frequencies f. An arbitrary amount of the signal energy can be restored in the Gabor features providing the reconstruction property. In addition, operations to perform deformation independent detection of events can be established [34].

1) Translation property: The first and most obvious use of the time-frequency representation is to detect time varying changes in a signal content. If the time axis is covered with Gabor functions and their overlap is sufficient, an event in an arbitrary position will always fall in the effective area of some of the functions. On this basis a translation invariant search can be established [34].

For the 1-D Gabor feature in (26) and a translated version

$$\xi_1(t) = \xi(t - t_1) \tag{27}$$

of some signal $\xi(t)$ it can be shown that [34]

$$r_{\xi_1}(t; f) = \int_{-\infty}^{\infty} \psi(t - t_\tau; f)\xi_1(t_\tau)dt_\tau = r_\xi(t - t_1; f) \tag{28}$$

which is the translation property of Gabor filters allowing a translation invariant detection of events. The translation property is common for most linear time-frequency representations and in general it holds for functions with a continuous translation parameter.

2) Scale property: While the translation property was recognized as an advantage already in the earliest studies of the short-time Fourier transform, the scale property did not gain any significant attention until the introduction of wavelets; from time-frequency to time-scale. The original time-frequency division by Gabor functions was a special case of the short-time Fourier transform and not comparable to time-scale methods [14]. By their mathematical properties both structures are special cases of "coherent states associated with a Lie group", but the short-time Fourier further corresponds to the Weyl-Heisenberg group while the wavelets correspond to the "ax+b" (affine) group [14], [23].

The time-frequency representation was criticized by Daubechies who showed distinct advantages for the affine group [14]. Fortunately, it is possible to formulate Gabor functions in an affine manner, as was established by connecting the sharpness α and frequency f in (9) [27]. The filters in (10) and (11) extract the same information in different scales, f denoting the scale parameter. Now, for a scaled signal $\xi_2(t) = \xi(at)$ where a denotes the scaling factor, it holds that ([27], [34])

$$r_{\xi_2}(t; f) = \int_{-\infty}^{\infty} \psi(t - t_\tau; f)\xi_2(t_\tau)dt_\tau = r_\xi(at; \frac{f}{a}) \ . \tag{29}$$

The interpretation of the above result is straightforward: a response for a signal is the same as the response of a similarly scaled filter for a scaled version of the signal. To maintain a homogenous spacing between the scales, a logarithmic frequency scale is typically established ([35], [14], [27], [34])

$$f_k = c^{-k}f_{max}, \text{ for } k = 0, \dots, m - 1 \tag{30}$$

where f_{max} is the maximum frequency and c is the frequency scaling factor. Some useful values for c include $c = 2$ for octave spacing and $c = \sqrt{2}$ for half-octave spacing. In Fig. 3

there is a logarithmic lattice marked with small circles and examples of correspondingly spaced Gabor filters.
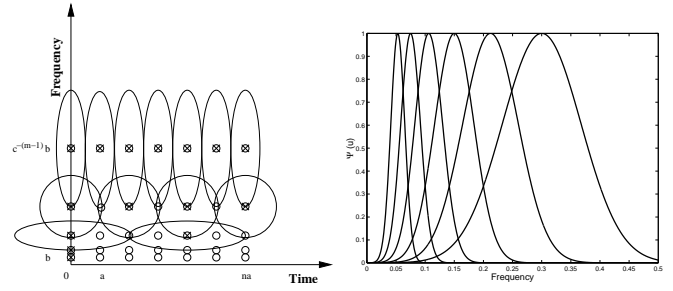

Fig. 3. Logarithmic lattice indicating the time-frequency spacing of Gabor filters $\psi(t - na; c^{-k}b)$. Only some examples of filter bandwidths (uncertainties) are shown.

It should be noted that filters on lower frequencies spread over a substantially larger time duration than filters on high frequencies, and thus, the logarithmic relation can be reflected also to the time domain (ellipses in Fig. 3) [14], [36].

B. Space–frequency features of 2-D signals

The results from previous section are next generalized to two dimensions since the feature extraction can be performed similarly in two dimensions via the convolution as

$$r_\xi(x, y; f, \theta) = \psi(x, y; f, \theta) * \xi(x, y)$$
$$= \int \int_{-\infty}^{\infty} \psi(x - x_\tau, y - y_\tau; f, \theta)\xi(x_\tau, y_\tau)dx_\tau dy_\tau \tag{31}$$

1) Physiology of vision: It seems that in 1-D signal processing the Gabor functions are mostly applied via the Gabor expansion, but in 2-D processing Gabor filtering has received much more attention. This can partly be explained by an active use of other time-frequency methods in 1-D. In 2-D one dominating reason has been correspondences found in the physiology of mammalian vision. The relevance of these considerations is not emphasized in this study but our reasoning is rather based on present results and established properties of Gabor filters. A comprehensive introduction to the mammalian vision can be found in [37].

2) Translation, scale, and rotation properties: The 1-D translation and scale properties in Eqs. (28) and (29) can be straightforwardly generalized to two dimensions [34], but there is a new variant, the orientation, which has no analogy with the 1-D case. The rotation of an object and filter responses have been studied in line detection [38], but were first introduced in a more general form by Würtz [32] and as a simple response matrix shift operation by the authors in [39], [34] and by others in [40].

A rotated version $\xi_3(x, y)$ of a 2-D signal $\xi(x, y)$, an image, rotated anti-clockwise around a spatial location (x_0, y_0) by an angle ϕ can be written as

$$\xi_3(x, y) = \xi(\hat{x}, \hat{y})$$
$$\hat{x} = (x - x_0)\cos\phi + (y - y_0)\sin\phi + x_0 \tag{32}$$
$$\hat{y} = -(x - x_0)\sin\phi + (y - y_0)\cos\phi + y_0$$

It can be easily shown that [34]

$$r_{\xi_3}(x_0, y_0; f, \theta) = r_\xi(x_0, y_0; f, \theta - \phi). \tag{33}$$

Finally, utilizing the translation, scale, and rotation properties of the 2-D Gabor features it can be concluded that for a 2-D signal $\xi'(x, y)$ which is translated from a location (x_0, y_0) to a location (x_1, y_1), scaled by a factor a and rotated anticlockwise by an angle ϕ around the location (x_1, y_1) it holds that

$$r_{\xi'}(x_1, y_1; f, \theta) = r_\xi(ax_0, ay_0; \frac{f}{a}, \theta - \phi) \tag{34}$$

which is a central result utilized in the rotation, scale, and translation invariant search of objects in images ([34], [32]).

3) Simple Gabor feature space: A simple feature space has been introduced by authors in [34] and successfully applied to face detection in [41], [42], [43]. A significant simplification made in the proposed feature space is the use of only one spatial location (x', y') to represent an object. The assumption is justified if the objects are simple or if they are distinguishable from each other in the feature space. This is not the case with, for example, the human face, but seems to hold between salient sub-parts, such as nostrils, eyes, mouth corners, etc. The filters in one location tuned to various frequencies and orientations span a sub-space whose accuracy decreases from the filter origin. This is demonstrated in Fig. 4 where an original facial image is reconstructed using filter responses from 10 locations with four orientations and five frequencies. The reconstruction from a few filters can be further improved by optimization [44].
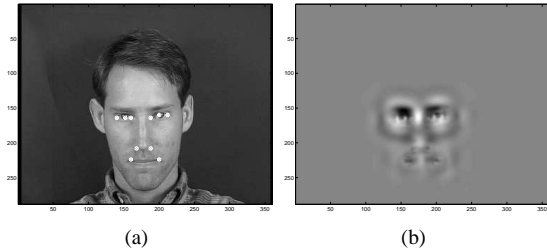

Fig. 4. Reconstruction from responses at 10 different locations: (a) original; (b) reconstruction.

Using a single location and the logarithmic frequency scale, the object search can be done independently in the scale and location, as the objects appear the same in the feature space [45], [34]. Furthermore, to search over orientations the filter orientations can be drawn from

$$\theta_k = \frac{k2\pi}{n}, \quad k = \{0, \dots, n-1\} \ . \tag{35}$$

Since for real signals the responses at $[\pi, 2\pi[$ are complex conjugates of responses on $[0, \pi[$ it is usual to compute responses only for the half plane

$$\theta_k = \frac{k\pi}{n}, \quad k = \{0, \dots, n-1\} \ . \tag{36}$$

Now, by using the filter responses in (31) at location (x_0, y_0) with the parameters drawn from Eqs. (30) and (36) a feature

matrix $\mathbf{G}$ can be constructed as [34]

$$\mathbf{G} = \begin{pmatrix} r(x_0,y_0;f_0,\theta_0) & r(x_0,y_0;f_0,\theta_1) & \cdots & r(x_0,y_0;f_0,\theta_{n-1}) \\ r(x_0,y_0;f_1,\theta_0) & r(x_0,y_0;f_1,\theta_1) & \cdots & r(x_0,y_0;f_1,\theta_{n-1}) \\ \vdots & \vdots & \ddots & \vdots \\ r(x_0,y_0;f_{m-1},\theta_0) & r(x_0,y_0;f_{m-1},\theta_1) & \cdots & r(x_0,y_0;f_{m-1},\theta_{n-1}) \end{pmatrix} \tag{37}$$

Operations for rotation and scale invariant searches of objects can be defined as a column-wise circular shift of the response matrix corresponding to the rotation of the object around the location (x_0, y_0) and a row-wise shift corresponding to the scaling of an object by a factor c [34]. An illumination invariance can be achieved by normalizing the feature matrix [34].

V. INVARIANT RECOGNITION

The translation, scale, and rotation properties provide operations to perform search over their variation, and thus, the object identification may concentrate only on discriminating objects in one pose. That is, features from objects are stored in a standard pose and the search is performed by applying invariance shifts of features in order to detect objects in arbitrary poses.

Various methods based directly on Gabor feature similarity measures have been used, for example the face detection and recognition methods in [46], [40], [47], [35]. An extensive work on proper similarity measures has been reported by Malsburg's research group [47], [35], [48], [49], [32]. The proposed similarity measures were originally based on disparity measures for binocular camera systems and are most useful when objects in images are roughly in the same scale and orientation. Lately, an ad hoc conditional likelihood response distribution based measure has been proposed [46].

The similarity value guides the search in the similarity based approaches, but in the simple Gabor feature space this is not necessary and an alternative, a classifier based recognition, can be used [41], [42], [43]. The classifier selection must be based on the behavior of the instances in the feature space. Due to the smooth shape of the Gabor filters the behavior of their complex responses is also typically smooth as demonstrated in Fig. 5 which shows responses for the real objects used in [41], [42], [43]. In Fig. 5 filter responses from left and right eye centers, which should be roughly symmetric, are plotted on two frequencies at zero orientation. Complex responses in Figs. 5(a) and 5(b) form smooth clusters, a result which can be used in the classifier selection, e.g., by applying a statistical classifier with Gaussian mixture-model probability densities [41], [42], [43]. Complex responses are most natural since, for example, the phase warping would cause a discontinuity in the case of a magnitude-phase pair (see Figs. 5(c) and 5(d)). In Fig. 5 it can be seen how the phase variation decreases at lower frequencies. In addition, the robustness and stability of the phase information increases for small values of sharpness parameters (wider bandwidth). It should be noted that the use of phase information is very important in recognition and detection of objects as has been practically demonstrated in [50] and especially in [51] where the information richness of the phase is analyzed using simple binarization of the phase.
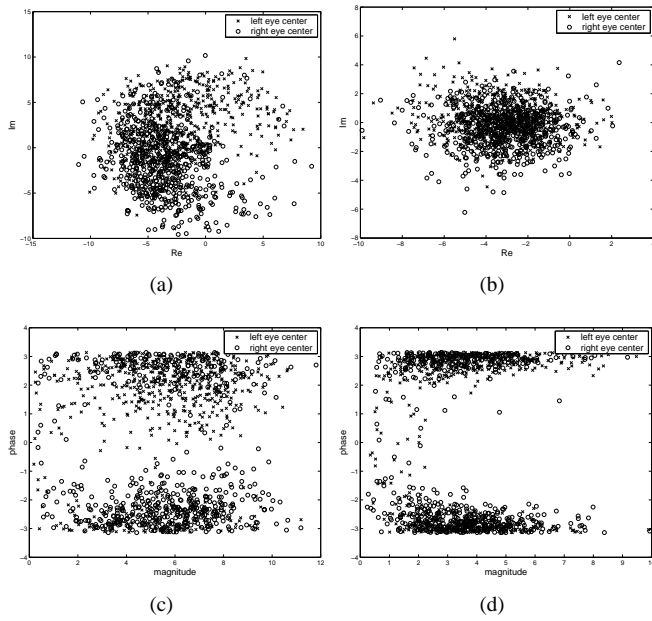
Fig. 5. Gabor features of the left and right eye centers (XM2VTS database) [41]: real-imaginary plots on (a) high frequency ($f = \frac{1}{3}$, $\theta = 0$); (b) low frequency ($f = \frac{1}{12}$, $\theta = 0$); magnitude-phase plots on (c) high frequency ($f = \frac{1}{3}$, $\theta = 0$); (d) low frequency ($f = \frac{1}{12}$, $\theta = 0$).

A. Practical considerations

There are many practical requirements to enable accurate and reliable use of Gabor features and they will briefly be discussed next. A more detailed discussion can be found in [25].

1) General: The sampling and quantization of Gabor filters was already discussed in Section III and restrictions were given for accurate generation of the discrete filters. In feature extraction, standard computer arithmetics are usually available, the filters are roughly of the same order of magnitude, and the distinguishing properties in well-posed problems should not be merely very small differences in filter responses. Thus the relevance of factors other than sampling and quantization is typically significant.

The implementation of Gabor feature extraction is an important computational consideration. For example, in the simple Gabor feature space a full search in a two dimensional image needs at least Image width × Image height × Row (scale) shifts × Column (orientation) shifts feature vectors to be classified. In practice heuristics and a priori information must be incorporated in order to reduce the computing time; objects can be searched only over a few orientations and scales, and not in every location [41], [42], [43]. These procedures are also motivated by the multi-resolution structure.

Typically cyclic convolution is assumed in the computations and one has to be aware that this affects the filter responses making them unreliable near signal ends or image edges. To have a reliable response at a location (x, y) the effective area of the filter envelope should be inside the image at that point. The minimum filter size, $L_{min} = L$, can be resolved from Eq. (23) or (25) and the responses within a distance $\frac{L_{min}}{2}$ from edges should be neglected.

Another important consideration is the selection of values for the frequency f, orientation θ, and sharpness γ and η. The selection of these values is the main problem in utilizing Gabor filters in feature extraction since they define the time-frequency division in the extraction. While some adjustments can be made for the frequency and orientation, the sharpness parameters γ and η are completely application dependent. As a rule of thumb, the filter size should be adjusted to encapsulate optimally the event of interest, in time and frequency. For example, in texture segmentation a circular form of the filter ($\gamma = \eta$) is typically used, but to distinguish objects of different dimensions a not equal aspect ratio ($\eta < \gamma$) may provide more accurate results [52]. On the other hand, selection of one parameter can be compensated in the selection of another [45]. This vague nature in the selection of the parameters is one of the reasons that an explicit optimization of the parameters has been applied [53], [54], but for some applications this may be impossible or at least cause inconvenience because the optimization must be repeated in new situations. For greater flexibility it is possible to select the filter parameters to cover the whole parameter space within the limits of the computing time and resolving the importance of different filters is left to a learning system, a classifier. In practice, as many filters as possible within the limits of the computational resources should be used, to extract as fine details as possible.

2) Shiftability: For efficient computation and limited size representation of the features in (31), the features must be calculated only for a sufficiently small number of parameters. This is evident since the discretization of the spatial coordinates (x, y), frequency f, and orientation θ corresponds to a four dimensional feature space. There is a trade-off between the selection of the filter parameters and the continuity of the feature space. In other words, one would like to fill the feature space as efficiently as possible to decrease the computation, but at the same time the filters should overlap as much as possible to provide smooth behavior in situations where events fall somewhere between two filters. This trade-off is a known problem between orthogonal and non-orthogonal presentations since the continuous behavior, shiftability, can be achieved only by relaxing the orthogonality property [55]. In shiftable systems the response in any location can be reconstructed as a linear combination from the filter responses at fixed locations. If the shiftability condition holds over more than one parameter, the system is said to be joint shiftable [55]. It should be noted that Gabor filters do not meet the requirements of exact shiftability, although they have the very beneficial property of being well localized, which is desirable in low-level image processing. In order to decrease the orthogonality, the overlap of the filters must be increased. The overlap cannot be accomplished by sharpness parameters γ and η alone but the number of filters must be increased. The smooth behavior, approximate shiftability, is again a trade-off which is achieved at the expense of computational complexity. Approximate shiftability is currently an unexplored property of Gabor filters and spaces spanned by Gabor functions and only a few studies of the definition of approximate shiftability have been done [56], [57], [58].

3) Robustness and noise tolerance: Robustness and noise tolerance are essential considerations for practical applications. Robustness here means the stability of the responses when conditions, such as the filter parameters, are changed. There are many natural variations in the conditions of real applications which cannot be explicitly covered in the method development phase, and thus, the methods benefit if they are robust against these changes: a small change in the conditions induces only a small change in the system performance. Robustness against variation in the filter parameters is addressed in [59] and by the authors in [33] where the classification accuracy is experimentally shown to smoothly and continuously change proportional to the change of the parameter values. It can be concluded that the accuracy decreases gradually as the deviation from the optimal parameter values increases; a phenomenon which can be understood from the Gaussian shapes of Gabor filters in both domains [60].

Since some noise and distortions can also be described as variations of the filter parameters, it is assumed that the behavior of the Gabor features is smooth in the presence of noise and distortions. Furthermore, the Gabor filters are optimally joint localized in time and frequency, and thus, distortions and noise present in distinct locations, time or frequency, do not significantly interfere with the filter responses. Noise and distortion tolerance is natural to Gabor filters and it has been considered in [60] where the Gabor filter based features perform outstandingly well in the presence of noise and distortions, namely Gaussian, salt-and-pepper and pixel displacement noise, and illumination gradient distortion. Of course noise removing methods can be applied and distortions can be eliminated, but in cases where these are not known a priori the Gabor filters still provide tolerance of a high degree.

VI. EXAMPLES

A. Symbol recognition

This first example utilizes a globally computed sum of Gabor responses, corresponding to edge histograms over different orientations, and thus, provides no elegance or novelty as a feature extraction method. However, similar simple features are often used in image query applications and it successfully demonstrates the rotation property and noise robustness of Gabor features. The demonstrated features, global Gabor features, have originally been proposed by the authors in [61] and the features would have their use in databases of engineering drawings [60], [33].

In a global Gabor feature, the responses are summed over an image to form a global feature [39]

$$G(f, \theta) = \sum_x \sum_y |r_\xi(x, y; f, \theta)| \ . \qquad (38)$$

The feature can be considered as a histogram of the responses in (31) for different frequencies f and orientations θ. Since edges appear at high frequencies and lines on their fundamental frequencies, a single properly selected frequency f is needed to represent a histogram of lines in different orientations. Now, a feature vector can be defined as

$$\mathbf{G} = (G(f, \theta_0) \ G(f, \theta_1) \ \dots \ G(f, \theta_{n-1})) \ . \qquad (39)$$

The global Gabor feature in (38) is translation invariant as the responses are summed over the whole image. For higher frequencies representing edges, the object scale affects response magnitudes, but the ratio between the magnitudes remains over different scales. Thus, a scale invariant feature can be constructed by normalizing the feature vector. In addition, the normalization makes the feature also illumination invariant, that is, invariant to a constant multiplier. Now, using the column-wise shift operation a rotation invariant similarity measure can be defined [34], [39]

$$d(\mathbf{G}_1, \mathbf{G}_2) = \min_k \{ \sum_{\theta_i} \left[\mathbf{G}_1(f, \theta_i) - \mathbf{G}_2^{(\theta+k)}(f, \theta_i) \right]^2 \} \quad (40)$$

where $\mathbf{G}_1$ and $\mathbf{G}_2$ are feature vectors from two inspected images.

The symbol images in Fig. 6 were used as groundtruth and their randomly rotated variants were used as the test set. The experiment was carried out by measuring distances to all base classes with the rotation invariant similarity measure in (40).
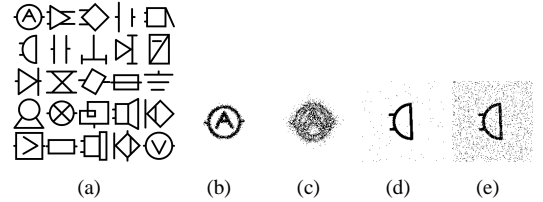

Fig. 6.　(a) symbols; (b) displacement noise ($\gamma_n = 1$); (c) ($\gamma_n = 0.2$); (d) salt-and-pepper noise ($\alpha_n = 0.4$); (e) ($\alpha_n = 2.0$).

The global Gabor features outperformed the previously used Hough transform based features [62] by being comparable in the presence of displacement noise and significantly more tolerant to salt-and-pepper noise (Fig. 7).
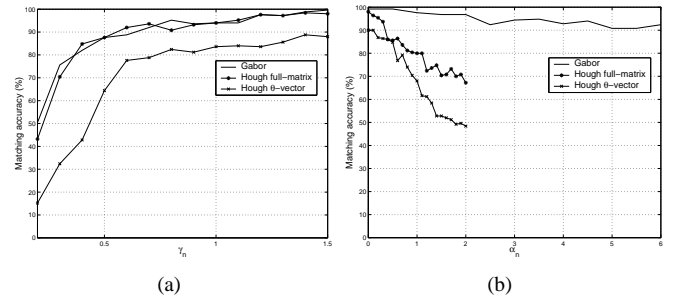

Fig. 7.　Matching accuracies as functions of noise parameters: (a) displacement noise; (b) salt-and-pepper noise. [61]

The results in Fig. 7 were achieved using the filter parameters $\gamma = 1$, $\eta = 1$, $f = 0.056$, and $N = 20$, but due to the smooth behavior of Gabor filter responses, the method is robust to a non-optimal selection of the filter parameters as demonstrated by the results in Fig. 8 ([60]).

The proposed method successfully utilizes the rotation invariance property of Gabor features and performs well also for a non-optimal selection of parameters. It should be noted that the size of discrete Gabor filters must be adequately large for accurate and reliable results. The inequalities in (24) and
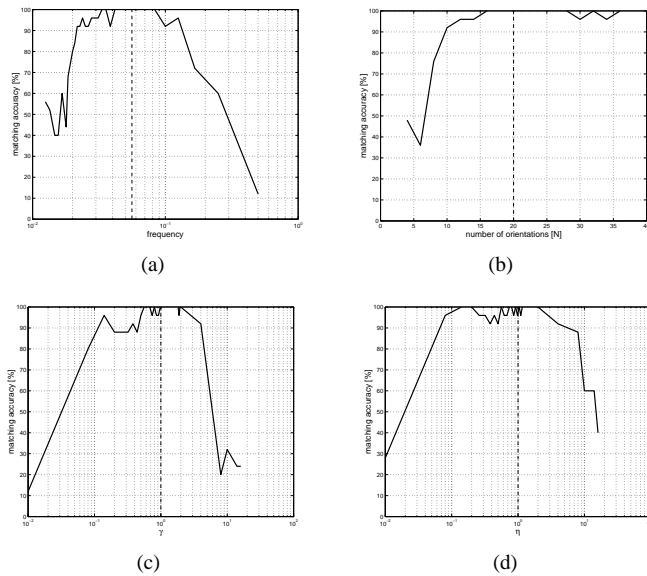
(a) (b)

Fig. 8. Matching accuracies as functions of the Gabor filter parameters (values used in Fig. 7 marked): (a) frequency f; (b) number of orientations N; (c) sharpness γ; (d) sharpness η.

(25) must not be violated and in practical computations the following values can be used as safe lower bounds $p_f = p_t = 0.9$.

B. Electric component detection

In this example, scale tuning of Gabor filters is used to detect objects of different sizes and rotation and translation properties are used to search objects over different locations and orientations.

The method is based on a detection of the fundamental frequencies of objects, that is, frequencies which describe the overall shape of an object [52], [60]. In a 1-D case a rectangle function may represent an ideal shape where the width of the rectangle is the size information to be inspected. It can be shown that the response of a Gabor filter at the centroid of the rectangle has a maximum on frequencies

$$f = \frac{1}{2w} + \frac{n}{w}, \quad n = 0, 1, 2, \dots \quad (41)$$

where w is the rectangle width [52]. The same result can be generalized to two dimensions. Now, to represent an object, e.g., an electrical component, frequencies with the maximum response are stored for all orientations describing the size variation of an object in different directions. To estimate also the pose of a detected object, the fundamental frequency features should be stored for objects in a standard pose.

It should be noted that the spacing of frequencies is not necessarily logarithmic but linear to achieve a sufficient resolution for size discrimination. The feature constructed by the method represents components by storing fundamental frequencies, which correspond to the size of an object along different orientations, as illustrated in Fig. 9. In the figure there is an object of width 80 and height 35 pixels when the corresponding fundamental frequencies are $\frac{1}{160}$ and $\frac{1}{70}$.
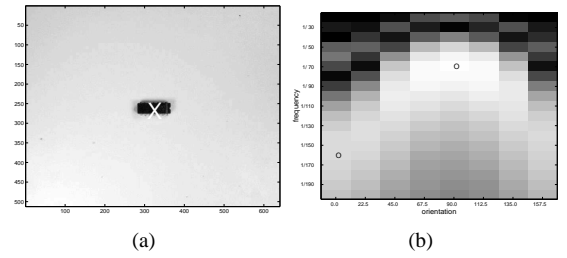

(a) (b)

Fig. 9. Response magnitudes: (a) component and its centroid ($(x, y) = (324, 263)$); (b) diagram of the filter responses at the centroid.

It should be noted that in the proposed method in [52], response matrix is not normalized until the frequencies with maximum responses have been located. However, in classification the responses are normalized to gain robustness against illumination changes. For this kind of problem the Fourier descriptors typically provide good results, but they need successful object segmentation which may become a difficult task in the presence of distortions and noise as in the example images in Fig. 10.
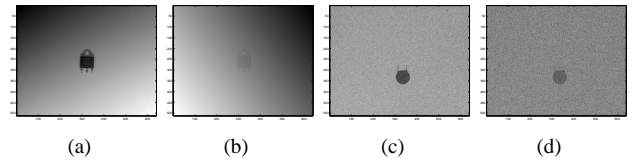

(a) (b) (c) (d)

Fig. 10. Noisy examples: (a) image gradient (slope 1.0); (b) (slope 2.5); (c) Gaussian noise (deviation 32); (d) (deviation 100).

The detection method in [52] does not need any preprocessing before the classification. To recognize eight different components at arbitrary poses in real images, a detection experiment was conducted using the fundamental frequency features and the rotation invariant similarity measure. There were 6 images of each component in random poses and random noise was applied to all of them repeatedly. The same experiment was also conducted using Hu's third order moments [63], but the results were significantly worse and with noise and distortions the moment invariant method completely failed. It is evident that the task is not trivial, yet the Gabor feature based method succeeded particularly well in the detection of electric components from noisy and distorted images and achieved 100% accuracy for images without distortions.

The proposed method provides a framework for simple object recognition tasks. Only one image from each class is needed for training and the only input information is the approximate centroids of the objects. After training, a simple algorithm can be used to detect and recognize stored objects from unseen images in a rotation and translation invariant manner and it is also possible to estimate the object pose. The achieved noise and distortion tolerances are evidence of the reliability the Gabor features may provide in practical applications.

C. Face evidence extraction

Recently, Hamouz and Kittler have proposed a complete framework for general object detection based on discriminative regions [64]. If an affine invariant detection of discriminative regions can be performed and spatial relationships between them are known, detection of an object can be done. The success of the framework depends on successful selection and efficient extraction of discriminative regions, e.g., facial features. In an application of the proposed method, the objects are frontal human faces and the discriminative regions are salient sub-parts, such as nostrils and eyes.

The detection of the discriminative regions, called facial evidences, was based on the simple Gabor feature space [34] where 10 evidence classes were inner and outer eye corners, eye centers, nostrils, and mouth corners on both sides of faces [42], [41]. The Gabor feature matrix responses were normalized to unity to achieve illumination invariance. Smooth behavior of features motivated the use of statistical classifiers: a sub-cluster classifier [41] and a Bayesian classifier using Gaussian mixture model [43]. The XM2VTS image database was used as a benchmark set [65]. In Fig. 11(a) is shown an example image and extracted evidence candidates which are passed to the next processing level. In Fig. 11(b) are shown the results of the face detection system as percents of images for which a specific error level was achieved. The error was defined as

$$d_{eye} = \frac{\max(d_l, d_r)}{\parallel C_l - C_r \parallel} \qquad (42)$$

where C_l and C_r are the correct eye center coordinates and d_l and d_r distances between the detected eye centers and the correct ones [66]. Translation, rotation, and scale invariant face
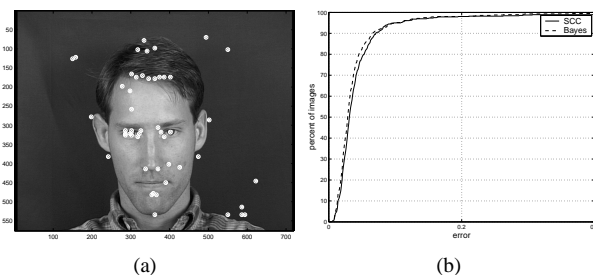

(a)　　　　　　(b)

Fig. 11. Face detection results: (a) examples of 5 best evidence candidates of the 10 classes. (b) face detection results for XM2VTS database using Gabor features with sub-cluster (SCC) and Bayesian classifiers.

evidence extraction was performed by shift operations in the simple Gabor feature space.

The results in Fig. 11 were achieved using the method in [42] and it outperforms the previous implementation of the same method with different features, Harris detectors [67] and also a competitive method by others [43], [66]. It is evident that full use of the invariance properties provides new application areas for Gabor features.

VII. Discussion

The motivation for this work was the success of Gabor features in many applications and problems in image processing.

It seems that one reason for their success and popularity is that changes in object location, scale, and orientation can be detected in the Gabor feature space. The background and use of the properties were presented in this study and they were demonstrated by application examples from literature.

The properties of Gabor filters seem to make invariant detection possible, but in addition, Gabor filters seem to establish a significant degree of robustness to photometric disturbances, such as illumination change and image noise, and to natural image variations, such as backgrounds.

There are new methods, wavelets being the most prominent, appearing in fields where time-frequency methods have traditionally been used and some may even claim that time has passed time-frequency representations by. However, in this study the invariance properties based on time-frequency information have been shown to be useful and new connections have been made to recently established concepts important in pattern recognition, such as shiftability, which is evidence that there is still unrevealed potential in feature extraction with Gabor filters. This development can also be seen in wavelet research where beneficial properties of redundant wavelets have been noted and the wavelets are being further developed towards a more general theory of frames. Future research will show if harmony between Gabor functions and wavelets can be established and whether a comprehensive framework for general efficient feature extraction can be realized utilizing properties of both; "the general picture processing operator" [7].

References

[1] D. Gabor, "Theory of communications," *Journal of Institution of Electrical Engineers*, vol. 93, pp. 429–457, 1946.

[2] H. Nyquist, "Certain factors affecting telegraph speed," *Bell System Technical Journal*, vol. 3, pp. 324–346, 1924.

[3] E. Shannon, "A mathematical theory of communication," *The Bell System Technical Journal*, vol. 27, pp. 397–423, 623–656, 1948.

[4] Martin J. Bastiaans, "Gabor's expansion of a signal into Gaussian elementary signals," *Proceedings of the IEEE*, vol. 68, no. 4, pp. 538–539, 1980.

[5] Martin J. Bastiaans, "Gabor's signal expansion and the Zak transform," *Applied Optics*, vol. 33, no. 23, pp. 5241–5255, 1994.

[6] Martin J. Bastiaans and Marc C.W. Geilen, "On the discrete Gabor transform and the discrete Zak transform," *Signal Processing*, vol. 49, no. 3, pp. 151–166, 1996.

[7] Goesta H. Granlund, "In search of a general picture processing operator," *Computer Graphics and Image Processing*, vol. 8, pp. 155–173, 1978.

[8] John G. Daugman, "Uncertainty relation for resolution in space, spatial frequency, and orientation optimized by two-dimensional visual cortical filters," *Journal of the Optical Society of America A*, vol. 2, no. 7, pp. 1160–1169, 1985.

[9] Tai Sing Lee, "Image representation using 2D Gabor wavelets," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 18, no. 10, pp. 959–971, 1996.

[10] Ronald N. Bracewell, *The Fourier Transform and its Applications*, McGraw-Hill, Inc., 2 edition, 1978.

[11] Frank Amoroso, "The bandwidth of digital data signals," *IEEE Communications Magazine*, vol. 18, no. 6, 1980.

[12] David Slepian, "On bandwidth," *Proceedings of the IEEE*, vol. 64, no. 3, pp. 292–300, 1976.

[13] A.H. Nuttall and F. Amoroso, "Minimum Gabor bandwidth of M orthogonal signals," *IEEE Transactions on Information Theory*, vol. 11, no. 3, pp. 440–444, 1965.

[14] Ingrid Daubechies, "The wavelet transform, time-frequency localization and signal analysis," *IEEE Transactions on Information Theory*, vol. 36, no. 5, pp. 961–1005, 1990.

[15] L. Wang, C.-T. Chen, and W.-C. Lin, "An efficient algorithm to compute the complete set of discrete Gabor coefficients," *IEEE Transactions on Image Processing*, vol. 3, no. 1, pp. 87–92, 1994.

[16] S. Li and Jr. M. Healy, "A parametric class of discrete Gabor expansions," *IEEE Transactions on Signal Processing*, vol. 44, no. 2, pp. 201–211, 1996.

[17] J. Yao, P. Krolak, and C. Steele, "The generalized Gabor transform," *IEEE Transactions on Image Processing*, vol. 4, no. 7, pp. 978–988, 1995.

[18] S. Qian and D. Chen, "Joint time-frequency analysis," *IEEE Signal Processing Magazine*, vol. 16, no. 2, pp. 52–67, 1999.

[19] B. Friedlander and B. Porat, "Detection of transient signals by the Gabor representation," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 37, no. 2, pp. 169–180, 1989.

[20] V.C. Chen and H. Ling, "Joint time-frequency analysis for radar signal and image processing," *IEEE Signal Processing Magazine*, vol. 16, no. 2, pp. 81–93, 1999.

[21] H. Bölcskei and F. Hlawatcsh, "Equivalence of DFT filter banks and Gabor expansion," in *SPIE Proceedings, Wavelet Applications in Signal and Image Processing III*, San Diego, CA, USA, 1995, vol. 2569, pp. 128–139.

[22] H. Bölcskei, F. Hlawatsch, and H.G. Feichtinger, "Frame-theoretic analysis of oversampled filter banks," *IEEE Transactions on Signal Processing*, vol. 46, no. 12, pp. 3256–3268, 1998.

[23] H.G. Feichtinger and T. Strohmer, Eds., *Gabor Analysis and Algorithms*, Birkhäuser, 1998.

[24] I. Daubechies, B. Han, A. Ron, and Z. Shen, "Framelets: MRA-based constructions of wavelet frames," *Applied and Computational Harmonic Analysis*, vol. 14, no. 1, pp. 1–46, 2003.

[25] Joni-Kristian Kämäräinen, *Feature Extraction Using Gabor Filters*, Ph.D. thesis, Lappeenranta University of Technology, 2003.

[26] Jont B. Allen, "Short term spectral analysis, synthesis, and modification by discrete Fourier transform," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 25, no. 3, pp. 235–238, 1977.

[27] A. Grossmann and J. Morlet, "Decomposition of Hardy functions into square integrable wavelets of constant shape," *SIAM Journal on Mathematical Analysis*, vol. 15, no. 4, pp. 723–736, 1984.

[28] L.J. Cutrona, E.N. Leith, C.J. Palermo, and L.J. Porcello, "Optical data processing and filtering systems," *IRE Transactions on Information Theory*, vol. 6, no. 3, pp. 386–400, 1960.

[29] David Casasent and Demetri Psaltis, "New optical transforms for pattern recognition," *Proceedings of the IEEE*, vol. 65, no. 1, pp. 77–78, Jan 1977.

[30] John G. Daugman, "Complete discrete 2-d Gabor transforms by neural networks for image analysis and compression," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 36, no. 7, pp. 1169–1179, 1988.

[31] Alan Conrad Bovik, Marianna Clark, and Wilson S. Geisler, "Multichannel texture analysis using localized spatial filters," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 12, no. 1, pp. 55–73, January 1990.

[32] Rolf P. Würtz, *Multilayer Dynamic Link Networks for Establishing Image Point Correspondences and Visual Object Recognition*, Ph.D. thesis, Ruhr-Universität Bochum, 1994.

[33] J.-K. Kamarainen, V. Kyrki, and H. Kälviäinen, "Robustness of Gabor feature parameter selection," in *Proceedings of the IAPR Workshop on Machine Vision Applications*, 2002, pp. 132–135.

[34] V. Kyrki, J.-K. Kamarainen, and H. Kälviäinen, "Simple Gabor feature space for invariant object recognition," *Pattern Recognition Letters*, vol. 25, no. 3, pp. 311–318, 2003.

[35] Joachim Buhmann, Martin Lades, and Christoph von der Malsburg, "Size and distortion invariant object recognition by hierarchical graph matching," in *IJCNN International Joint Conference on Neural Networks*, 1990, vol. 2, pp. 300–311.

[36] Stephane G. Mallat, "A theory for multiresolution signal decomposition: The wavelet representation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 11, no. 7, pp. 674–693, 1989.

[37] S.E. Palmer, *Vision Science - Photons to Phenomenology*, MIT Press, 1999.

[38] Jian Chen, Yoshinobu Sato, and Shinichi Tamura, "Orientation space filtering for multiple orientation line segmentation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, no. 5, pp. 417–429, May 2000.

[39] V. Kyrki, J.-K. Kamarainen, and H. Kälviäinen, "Invariant shape recognition using global Gabor features," in *12th Scandinavian Conference on Image Analysis*, Bergen, Norway, June 2001, pp. 671–678.

[40] Hyun Jin Park and Hyun Seung Yang, "Invariant object detection based on evidence accumulation and Gabor features," *Pattern Recognition Letters*, vol. 22, pp. 869–882, 2001.

[41] J.-K. Kamarainen, V. Kyrki, M. Hamouz, J. Kittler, and H. Kälviäinen, "Invariant Gabor features for face evidence extraction," in *Proceedings of the IAPR Workshop on Machine Vision Applications*, Nara, Japan, 2002, pp. 228–231.

[42] M. Hamouz, J. Kittler, J.-K. Kamarainen, and H. Kälviäinen, "Hypotheses-driven affine invariant localization of faces in verification systems," in *Proceedings of the International Conference on Audio- and Video-Based Biometric Person Authentication*, 2003, pp. 276–284.

[43] M. Hamouz, J. Kittler, J.-K. Kamarainen, P. Paalanen, and H. Kälviäinen, "Affine-invariant face detection and localization using GMM-based feature detector and enhanced appearance model," in *Proceedings of the 6th International Conference on Automatic Face and Gesture Recognition (to be published)*, 2004.

[44] V. Krüger and G. Sommer, "Gabor wavelet networks for efficient head pose estimation," *Image and Vision Computing*, vol. 20, no. 9-10, pp. 665–672, 2002.

[45] Moshe Porat and Yehoshua Y. Zeevi, "The generalized Gabor scheme of image representation in biological and machine vision," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 10, no. 4, pp. 452–468, 1988.

[46] J. Lampinen, T. Tamminen, T. Kostiainen, and I. Kalliomäki, "Bayesian object matching based on MCMC sampling and Gabor filters," in *Proceedings of the SPIE Intelligent Robots and Computer Vision XX: Algorithms, Techniques, and Active Vision*, 2001, vol. 4572, pp. 41–50.

[47] J. Buhmann, J. Lange, and C. Von Der Malsburg, "Distortion invariant object recognition by matching hierarchically labeled graphs," in *IJCNN: International Joint Conference on Neural Networks*, 1989, vol. 1, pp. 155–159.

[48] M. Lades, J.C. Vorbrüggen, J. Buhmann, J. Lange, C. von der Malsburg, R.P. Würtz, and W. Konen, "Distortion invariant object recognition in the dynamic link architecture," *IEEE Transactions on Computers*, vol. 42, no. 3, pp. 300–311, 1993.

[49] L. Wiskott, J.-M. Fellous, N. Krüger, and C. von der Malsburg, "Face recognition by elastic bunch graph matching," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 19, no. 7, pp. 775–779, 1997.

[50] A.V. Oppenheim and J.S. Lim, "The importance of phase in signals," *Proceedings of the IEEE*, vol. 69, no. 5, pp. 529–541, 1981.

[51] John Daugman, "Statistical richness of visual phase information: Update on recognizing persons by iris patterns," *International Journal of Computer Vision*, vol. 45, no. 1, pp. 25–38, 2001.

[52] J.-K. Kamarainen, V. Kyrki, and H. Kälviäinen, "Fundamental frequency Gabor filters for object recognition," in *Proceedings of the 16th International Conference on Pattern Recognition*, Quebec, Canada, 2002, vol. 1, pp. 628–631.

[53] D. Dunn and W.E. Higgins, "Optimal Gabor filters for texture segmentation," *IEEE Transactions on Image Processing*, vol. 4, no. 7, pp. 947–964, 1995.

[54] Anil K. Jain and F. Farrokhnia, "Unsupervised texture segmentation using Gabor filters," *Pattern Recognition*, vol. 24, no. 12, pp. 1167–1186, 1991.

[55] Eero P. Simoncelli, William T. Freeman, Edward H. Adelson, and David J. Heeger, "Shiftable multiscale transforms," *IEEE Transactions on Information Theory*, vol. 38, no. 2, pp. 587–607, Mar. 1992.

[56] J. Sampo, J.-K. Kamarainen, M. Heiliö, and H. Kälviäinen, "Measuring shiftability of frames of regular translates," in *Proceedings of the Nordic Signal Processing Conference*, 2004, To be published.

[57] Qiao Wang, "The shiftability of some wavelet bases," *Computers and Mathematics with Applications*, vol. 40, pp. 957–964, 2000.

[58] Qiao Wang, "Classification of wavelet bases by translation subgroups and nonharmonic wavelet bases," *Acta Mathematica Sinica. English Series*, vol. 16, no. 2, pp. 307–312, 2000.

[59] R. Mehrotra, K.R. Namuduri, and N. Ranganathan, "Gabor filter–based edge detection," *Pattern Recognition*, vol. 25, no. 12, pp. 1479–1494, 1992.

[60] J.-K. Kamarainen, V. Kyrki, and H. Kälviäinen, "Noise tolerant object recognition using Gabor filtering," in *Proceedings of the 14th International Conference on Digital Signal Processing*, Santorini, Greece, 2002, vol. 2, pp. 1349–1352.

[61] V. Kyrki, J.-K. Kamarainen, and H. Kälviäinen, "Content-based image matching using Gabor filtering," in *Proceedings of the International Conference on Advanced Concepts for Intelligent Vision Systems Theory and Applications*, Baden-Baden, Germany, 2001, pp. 45–49.

[62] Pasi Fränti, Alexey Mednonogov, V. Kyrki, and H. Kälviäinen, "Content-based mathing of line-drawing images using the Hough transform," *International Journal on Document Analysis and Recognition*, vol. 3, pp. 117–124, 2000.

[63] S.O. Belkasim, M. Shridhar, and M. Ahmadi, "Pattern recognition with moment invariants: A comparative study and new results," *Pattern Recognition*, vol. 24, no. 12, pp. 1117–1138, 1991.

[64] M. Hamouz, J. Kittler, J. Matas, and P. Bílek, "Face detection by learned affine correspondences," in *Joint international Workshop on Syntactical & Structural & Statistical Pattern Recognition*, Windsor (Ontario), Canada, August 2002, pp. 566–575.

[65] K. Messer, J. Matas, J. Kittler, J. Luettin, and G. Maitre, "XM2VTSDB: The extended M2VTS Database," in *Proceedings of Second International Conference on Audio and Video-based Biometric Person Authentication*, R. Chellapa, Ed., 1999, pp. 72–77.

[66] O. Jesorsky, K.J. Kirchberg, and R.W. Frischholz, "Robust face detection using the Hausdorff distance," in *Proceedings of the Third International Conference on Audio- and Video-based Biometric Person Authentication*, 2001, pp. 90–95.

[67] J. Matas, P. Bílek, M. Hamouz, and J. Kittler, "Discriminative regions for human face detection," in *Proceedings of ACCV2002: The 5th Asian Conference on Computer Vision*, 2002.