

Frequency map by Structure Tensor in Logarithmic Scale Space and Forensic Fingerprints

Josef Bigun
Halmstad University
30118 Halmstad, Sweden
josef.bigun@hh.se

Anna Mikaelyan
Halmstad University
30118 Halmstad, Sweden
anna.mikaelyan@hh.se

Abstract

Increasingly, absolute frequency and orientation maps are needed, e.g. for forensics. We introduce a non-linear scale space via the logarithm of trace of the Structure Tensor. Therein, frequency estimation becomes an orientation estimation problem. We show that this offers significant advantages, including construction of efficient isotropic estimations of dense maps of frequency. In fingerprints, both maps are shown to improve each other in an enhancement scheme via Gabor filtering. We suggest a novel continuous ridge counting method, relying only on dense absolute frequency and orientation maps, without ridge detection, thinning, etc. Furthermore, we present new evidence that frequency maps are useful attributes of minutiae. We verify that the suggested method compares favorably with state of the art using forensic fingerprints as test bed, and test images where the ground truth is known. In evaluations, we use public data sets and published methods only.

1. Introduction

We study dense maps of absolute frequency, where *absolute* refers to the norm of a frequency vector. However, for brevity, the term frequency will be used here. We aim to estimate frequency maps for continuous ridge counting, and as minutia descriptors. The state of the art and contributions are given after problem statements below.

Theoretical problem

Conceptually, we have an ideal model of a family of functions $f \in \mathbb{R}$, defined on a dense coordinate set $\mathbf{r} \in \mathbb{R}^n$,

$$f(\mathbf{r}) = \sum_{k \in Q} A_k \cos(\boldsymbol{\omega}_k^T \mathbf{r} + \varphi_k) \quad (1)$$

where $A_k, \varphi_k \in \mathbb{R}$, are amplitudes and phases of sinusoids, respectively. It is assumed that $A_k > 0$ because signs can be absorbed by φ_k . Furthermore, the frequency vectors $\boldsymbol{\omega}_k$ share the same frequency¹, $\omega_0 = \|\boldsymbol{\omega}_k\| > 0$.

¹In contrast to $\boldsymbol{\omega}_k$, the subscript of ω_0 is a label (not index).

The set Q is finite and consists of the integers, $1, 2, \dots, M$. The frequency vectors are in the same Fourier domain half, and $0 < \boldsymbol{\omega}_k^T \boldsymbol{\omega}_k / \omega_0^2 \leq \delta$ i.e. no two vectors have an angle smaller than $\arccos(\delta)$ between. The latter is demanded to discern sinusoids w.r.t. their orientations.

We wish to estimate ω_0 which best fits the model (1) to a local image without knowing (the direction of) $\boldsymbol{\omega}_k$ nor introducing measurement artifacts depending on input orientation, including orientation bias. No-orientation artifacts in estimations is a desirable property in applications and called here isotropy. The opposite, unisotropy, is to be avoided. Example unisotropy occurs in edge magnitude extraction if horizontal and vertical edges/sinusoids have higher magnitudes than those of other directions systematically. We wish thus to have the same level of accuracy in ridge frequency estimation regardless the directions of input sinusoids/edges to be measured. We focus on fingerprints having $n = 2$, but the method can be adopted for other uses and $2 < n$. affect critical decisions.

Application problem

Fingermarks (or *latents*) are traces of fingers on objects at crime scenes, usually producing poor quality images. *Tenprints* are high quality fingerprints imaged, e.g. at country borders and embassies, Fig. 3 and 6.

It is desirable to support fingerprint examiners, jury, lawyers, and judges by automatically extracted reliable measurements but these should be verifiable and understandable at their discretion. The question is if automatically extracted i) ridge counting along any path, and ii) frequency maps have added value for evidence in identification. This work assumes the current legal procedures of decision making is in place, i.e. that the minutiae in a fingerprint will be extracted by a human expert whereas minutia extracton in tenprints will be fully automatic. Additionally, we assume that the human expert will be assisted by a machine for suggesting additional features (ridge-counts) along paths defined by minutiae that the expert has extracted.

State of the art

Witkin [18], have suggested multi-scale representations of images by successive Gaussian filterings. Lindeberg 1998, [14] proposed a general scale selection principle to estimate the local scale. This is evaluated further in Section 4.

Local orientation characteristics based theories have increasingly come to dominate mammalian vision [8] and local description in computational vision. A principle tool of modeling and extracting local orientation has been the Structure Tensor, derived as a spectral optimization in Bigun and Granlund [1], with well-evaluated isotropy properties. Isotropy and accuracy are critically important to many applications but especially so for decisions in forensic fingerprints. The method has been producing nearly all dense orientation maps in fingerprints and in numerous other applications. Structure Tensor can also be formulated as an extension of quadrature filtering as in Knutsson [12]. The quotients of quadrature filters can be used to estimate the local frequency [13], and will be evaluated here.

The usefulness of Structure Tensor has also been observed by di Zenzo and Silvano [3], Förstner and Gülch [5], Kass and Witkin [11]. The Generalized Structure Tensor has been an alternative to generalized Hough transform [7] to recognize arbitrary shape, including fingerprint cores and deltas, [17].

In the context of tenprints, [6], and [15] suggested a method of frequency estimation when orientations of local images are known. The orientations are computed from the eigen-vectors of the Structure Tensor [1]. As will be shown, the assumption of unique orientation (computed [6], or not [13]) is not a necessary condition, and may undermine frequency estimations if it does not hold, Sections 4, 7.

The dense orientation and frequency maps of [2] were estimated by applying the Fourier Transform to 32x32 (overlapping) image blocks. These were used to reject outliers and to reinforce one frequency and one orientation per image block in palmprints, [10]. The orientation estimator of [2], and that of the Structure Tensor can model the orientation as weighted averages of power spectra, the weights being similar $\omega^2 \exp 2i\angle\omega$ (eq. (8) of [2] vs Lemma 1, and eqs. (14-16) of [1]). This hints that both maps are conceptually from the same estimator, one realized in the spatial and the other in the Fourier domain, motivating fundamental studies of dense maps to complement the understanding of the practice.

Image based features have the potential to improve the recognition of fingermarks compared to using minutia locations and directions (constellations), [9]. However, in the latter image based features were computed by commercial, unpublished software and on tenprints. This suggests further studies of feature extraction on tenprints and fingermarks with published algorithms for better understanding the potentials and limitations.

Contributions and organization

We summarize first the Structure Tensor in Section 2. In itself this is not new, but the summary is motivated in that the summary is used to build up the novelties.

An initial contribution is to elucidate how Structure Tensor evolves isotropically as its “scale” changes. The derivation is given in Section 3.1, and 3.2.

It will be shown that frequency estimation is equivalent to orientation estimation in the scale space obtained by the logarithm of trace of Structure Tensor, Section 3.3. This novelty offers a mathematical unification of orientation and frequency estimation. A practical consequence is that with few scale-space samples frequency maps along with their certainties based on modelling errors, can be obtained.

In Section 4, we evaluate the method when the ground truth is known by using planar waves and noise. The results provide evidence for that our estimator is more isotropic and accurate w.r.t. the ground truth, in comparison to state of the art. We think that this is novel.

An iterative scheme to improve frequency maps of poor quality images is introduced, Section 6. Using Gabor filters, the scheme offers improved dense orientation maps too. Both maps are evaluated in comparative tests using fingermarks, Section 7.

We introduce an original use of frequency maps, to count ridges continuously in fingerprints along arbitrary paths, without detecting ridges, Section 5. The method allows automatic ridge counting in forensic fingermarks, at a quality not demonstrated before (83 % reliability), Section 7.

2. Orientation and Structure Tensor

Structure Tensor is a symmetric positive semi-definite matrix \mathcal{S} . It summarizes the statistics of directional contents of a local image $f(\mathbf{r})$, with $\mathbf{r} \in \mathbb{R}^n$

$$\mathcal{S} = \int_{\mathbf{r} \in \mathbb{R}^n} \nabla f(\mathbf{r}) \cdot \nabla^T f(\mathbf{r}) d\mathbf{r} \quad (2)$$

where f represents brightness (gray-value). Here, and in the sequel the integration domain is $\mathbf{r} \in \mathbb{R}^n$. However, for convenience of reading and compactness, this fact and $d\mathbf{r}$ are omitted from integrals in the sequel.

In 2D, a linear combination of the elements of \mathcal{S} directly yields the direction of the most significant eigenvector along with differences of eigenvalues, readily expressed as a second order complex moment (of the power spectrum).

$$I_{20} = \int (D_x f)^2 - \int (D_y f)^2 + i \cdot 2 \int (D_x f)(D_y f) \quad (3)$$

Then, the most significant eigenvector of \mathcal{S} , \mathbf{u}_{\max} , and the eigenvalues of \mathcal{S} , λ_{\max} and λ_{\min} are connected to I_{20} via:

$$\angle I_{20} = 2\angle \mathbf{u}_{\max}, \quad |I_{20}| = \lambda_{\max} - \lambda_{\min} \quad (4)$$

The trace of \mathbf{S} , $\text{Tr}(\mathbf{S})$, is the other second order complex moment (although it is always real valued),

$$I_{11} = \int (D_x f)^2 + \int (D_y f)^2 = \lambda_{\max} + \lambda_{\min}, \quad |I_{20}| \leq I_{11} \quad (5)$$

The inequality is a consequence of both triangle and Schwartz inequalities. It is tight, i.e. it holds with equality if and only if a single orientation fits the image perfectly (linear symmetry), [1]. However, the inequality is meaningful only in 2D as I_{20} is defined only in 2D, (3). By contrast, because I_{11} is real and equals to $\text{Tr}(\mathbf{S})$ in 2D, we define it as $I_{11} = \text{Tr}(\mathbf{S})$ for nD below, making I_{11} well defined even for $2 < n$.

Nonetheless, I_{20} and the inequality will be useful. This is because it will be shown that frequency estimation is equivalent to line-fitting in the $2D$ space of $\text{Tr}(\mathbf{S})$, and σ^2 (scale), which is a $2D$ orientation estimation problem, regardless n , the variable fixing the dimension of \mathbf{r} , and \mathbf{S} .

Defining a Gaussian g_n in nD with variance σ^2 as

$$g_n(\mathbf{r}, \sigma^2) = \frac{1}{(2\pi\sigma^2)^{n/2}} \exp\left(-\frac{\|\mathbf{r}\|^2}{2\sigma^2}\right), \quad (6)$$

the steps of dense orientation map construction by Structure Tensor are as follows.

- i) Convolve the original (large) image f to obtain the gradient image

$$\nabla f = h * f \quad (7)$$

where h is obtained by sampling the gradient of a Gaussian g_n (6), containing the Gaussian as factor,

$$h(\mathbf{r}, \sigma_{in}^2) = \nabla g_n(\mathbf{r}, \sigma_{in}^2) = -\frac{\mathbf{r}}{\sigma_{in}^2} g_n(\mathbf{r}, \sigma_{in}^2) \quad (8)$$

with σ_{in}^2 fixing its inner-scale, the frequency range retained by (7). This corresponds to $D_x f + iD_y f$ in 2D.

- ii) Apply pointwise tensor product to obtain

$$\tilde{\mathbf{S}} = \nabla f \nabla^T f \quad (9)$$

which represents \mathbf{S} in “zero-sized” neighborhoods (infinitesimal linear symmetry). This corresponds to $(D_x f + iD_y f)^2$ in 2D.

- iii) Convolve $\tilde{\mathbf{S}}$ with a Gaussian filter $g_n(\mathbf{r}, \sigma_{out}^2)$ defining the local image (the outer-scale).

$$\mathbf{S} = g_n(\mathbf{r}, \sigma_{out}^2) * \tilde{\mathbf{S}} \quad (10)$$

To obtain dense maps in 2D with complex pixels representing orientation vectors, the complex image $(D_x f + iD_y f)^2$ is convolved by g_n to yield I_{20} , (3), whereas I_{11} , the local upper bound of $|I_{20}|$, assessing local orientation estimation quality, is obtained by convolving the absolute values $|D_x f + iD_y f|^2$ with g_n , (5).

3. Frequency estimation

3.1. Fourier Transform and gradient

Gradient filtering, (7), is needed to obtain (local) Structure Tensor and the filter is given by, (8). Symbolized by H , FT of components of h are essentially mapped on h :

$$H(\boldsymbol{\omega}, \sigma_{in}^{-2}) = i\boldsymbol{\omega} \exp\left(-\frac{\|\boldsymbol{\omega}\|^2}{2\sigma_{in}^{-2}}\right) = i\boldsymbol{\omega} G_n(\boldsymbol{\omega}, \sigma_{in}^{-2}) \quad (11)$$

where $\boldsymbol{\omega} \in \mathbb{R}^n$ and G_n is FT of g_n (6), a Gaussian in essence

$$G_n(\boldsymbol{\omega}, \sigma_{in}^{-2}) = \exp\left(-\frac{\|\boldsymbol{\omega}\|^2}{2\sigma_{in}^{-2}}\right) \quad (12)$$

such that its “variance” is σ_{in}^{-2} and its value at the origin is 1. Once σ_{in}^2 is fixed, G_n depends on the length $\omega = \|\boldsymbol{\omega}\|$, i.e. it is $1D$ in ω even if $\boldsymbol{\omega}$ is nD . Since it is independent of direction of $\boldsymbol{\omega}$, G_n is isotropic. Thereby $\|H\|$ (but not H)

$$\|H(\boldsymbol{\omega}, \sigma_{in}^{-2})\| = \omega \exp\left(-\frac{\omega^2}{2\sigma_{in}^{-2}}\right) = \omega G_1(\omega, \sigma_{in}^{-2}) \quad (13)$$

where $\omega = \|\boldsymbol{\omega}\|$, is isotropic. It attains its maximum at all nD points on a ball with radius $\|\boldsymbol{\omega}\| = \hat{\omega}$

$$\hat{\omega} = \underset{\omega}{\text{argmax}} \|H\| = \frac{1}{\sigma_{in}} \quad (14)$$

3.2. Isotropic, spatially invariant energy by Trace

We assume, for now, that the input is a planar wave with a unique frequency vector, $\boldsymbol{\omega}_1$

$$f = A \cos(\boldsymbol{\omega}_1^T \mathbf{r}) \quad (15)$$

that is, $0 < A_1 = A$ and $\|\boldsymbol{\omega}_1\| = \omega_0$ and $A_k = 0$ for $k > 1$ in (1). The s presumption will be widened to include M sinusoids and will be stated as a theorem later, the present restriction serving as a pedagogical tool illustrating the signal processing mechanism behind our frequency estimation.

Using (11) we obtain the result of Step i) above as,

$$h * f(\mathbf{r}) = -\boldsymbol{\omega}_1 G_n(\boldsymbol{\omega}_1, \sigma_{in}^{-2}) A \sin(\boldsymbol{\omega}_1^T \mathbf{r}) \quad (16)$$

As it stands, (16) has significant spatial variation, due to the amplified sinusoid changing with \mathbf{r} . Thus, (16) is not suitable to estimate $\|\boldsymbol{\omega}_1\|$ densely and in a spatially invariant manner. The latter is due because $\boldsymbol{\omega}_1$ is constant w.r.t. \mathbf{r} whereas $h * f$ changes as \mathbf{r} does. Using the property inherent to the Structure Tensor, doubling frequency content of the input followed by low-pass filtering, essentially ripple-free (spatially invariant) estimations of $\|A \cdot H(\boldsymbol{\omega}_1, \sigma_{in}^{-2})\|^2$ can be obtained via the trace I_{11} , as below.

Following up the input of (15), the trace is computed in a Gaussian window having the variance σ_{out}^2 ,

$$I_{11}(\mathbf{r}', \sigma_{in}^2, \sigma_{out}^2) = \int \|h * f(\mathbf{r} + \mathbf{r}')\|^2 g_n(\mathbf{r}, \sigma_{out}^2) \quad (17)$$

$$= (A \cdot \omega_0 G_n(\omega_1, \sigma_{in}^{-2}))^2 \int \sin^2(\omega_1^T(\mathbf{r} + \mathbf{r}')) g_n(\mathbf{r}, \sigma_{out}^2)$$

where we can substitute the quadratic sinusoid with its equivalent having the double frequency

$$I_{11}(\cdot) = \|A \cdot H\|_2^2 \frac{1}{2} [1 - \int \cos(2\omega_1^T(\mathbf{r} + \mathbf{r}')) g_n(\mathbf{r}, \sigma_{out}^2)] \quad (18)$$

The integral term is an ordinary convolution of a sinusoid with the Gaussian window representing the outer scale. The convolution of a sinusoid yields also a sinusoid so that

$$I_{11}(\cdot) = \|AH(\omega_1, \sigma_{in}^{-2})\|_2^2 \frac{1}{2} [1 - G_n(2\omega_1, \sigma_{out}^{-2}) \cos(2\omega_1^T \mathbf{r}')] \quad (19)$$

where the ripple amplitude is $G_n(2\omega_1, \sigma_{out}^{-2})$.

The vacillation amplitude is determined by $2\omega_0$, exciting the Gaussian insignificantly, compared to ω_0 . Assuming that a rough range of ω_0 is known, $\omega_0 \in [\omega_{\min}, \omega_{\max}]$ (or a bandpass filtering is applied to f), the ripple amplitude

$$G_n(2\omega_1, \sigma_{out}^{-2}) = G_1(2\omega_0, \sigma_{out}^{-2}) = \exp(-\frac{(2\omega_0)^2}{2\sigma_{out}^{-2}}) \leq \epsilon$$

can be contained. The inequality can be rewritten as

$$\frac{1}{2\omega_0^2} \log\left(\frac{1}{\epsilon}\right) \leq \sigma_{out}^2 \quad (20)$$

For example, for $G_n(2\omega_1, \sigma_{out}^2) < 0.007$, i.e. $\epsilon = 0.007$, we obtain via (20) that σ_{out} must be $1.575\omega_{\min}^{-1}$, or larger. The response energy, $\|A \cdot H\|^2$ with prescribed accuracy can be estimated via I_{11} according to (17), where σ_{out} is fixed by the range of ω_0 , via (20).

$$I_{11}(\mathbf{r}', \sigma_{in}^2) \approx \frac{1}{2} \|AH(\omega_1, \sigma_{in}^{-2})\|^2 = \frac{A^2 \omega_0^2}{2} G_1^2(\omega_0, \sigma_{in}^{-2}) \quad (21)$$

We have not σ_{out}^2 as argument of I_{11} since it is now fixed.

The next theorem, the proof of which is similar to the procedure above and omitted for the convenience of reading in limited space, essentially states that Equation (21) holds even if the input consists of M sinusoids, (1). Its evaluation will however be given experimentally e.g. when the input consists in two sinusoids, in Section 4.

Theorem (Trace constancy). *The Trace of Structure Tensor, I_{11} , with parameters σ_{in} , σ_{out} obtained in an image region obeying the same single frequency model of (1) is nearly constant in the region, i.e. it differs absolutely from a constant with the prescribed amount of $\epsilon_1 \in \mathbb{R}$,*

$$|I_{11}(\mathbf{r}, \sigma_{in}^2) - \frac{A^2}{2} \omega_0^2 \exp(-\frac{\omega_0^2}{\sigma_{in}^{-2}})| \leq \epsilon_1 \quad \text{with } \mathbf{r} \in \mathbb{R}^n \quad (22)$$

where $0 < \omega_0$ and $A = \sqrt{A_1^2 + \dots + A_M^2}$, provided that σ_{out} is sufficiently large compared to ω_0^{-1} of the input.

3.3. Implementation

Using the Theorem above, (22), we can obtain ω_0 by applying logarithm to I_{20} and to its model, $\frac{A^2}{2} \omega_0^2 \exp(-\frac{\omega_0^2}{\sigma_{in}^{-2}})$,

$$\log I_{11}(\mathbf{r}', \sigma_{in}^2) = \log\left(\frac{1}{2} A^2 \omega_0^2\right) - \omega_0^2 \sigma_{in}^2 \quad (23)$$

which proves the next theorem, valid for any dimension n .

Theorem (Log-trace in scale-space). *Logarithm of the Structure Tensor trace constructs a non-linear scale space in which signals fulfilling (1) generate a line*

$$\log I_{11}(\mathbf{r}, \sigma_{in}^2) = C_A - \omega_0^2 \sigma_{in}^2 \quad \text{with } \mathbf{r} \in \mathbb{R}^n \quad (24)$$

with direction determined by $-\omega_0^2$, and constant $C_A = \log(2^{-1} A^2 \omega_0^2)$ being independent of the scale σ_{in}^2 .

Tangents of the parametric curve generated by s

$$\mathbf{s}(\tau) = (\sigma_{in}^2(\tau), \log I_{11}(\mathbf{r}, \sigma_{in}^2(\tau))) \quad (25)$$

where τ is the parameter, determine $\omega_0(\tau)$. If and only if the curve is a line, ω_0 is constant, i.e. tangent vectors are parallel. This scale space is always two dimensional even if the dimension of \mathbf{r} is not.

The 2D tensor fits Total Least Square, TLS, angles (3,5),

$$I_{20}^s = \int_{\tau \in [\tau_{\min}, \tau_{\max}]} \left(\frac{d\sigma_{in}^2(\tau)}{d\tau} + i \frac{d \log I_{11}}{d\tau} \right)^2 d\tau \quad (26)$$

$$I_{11}^s = \int_{\tau \in [\tau_{\min}, \tau_{\max}]} \left(\frac{d\sigma_{in}^2(\tau)}{d\tau} \right)^2 + \int \left(\frac{d \log I_{11}}{d\tau} \right)^2 d\tau \quad (27)$$

with s referring to the logarithmic scale space (not the image space), and with integrations over a range of τ determined next. Equations (26)-(27) deliver the double of TLS tangent angle in the argument of I_{20}^s , and the error sum. Then, ω_0 is estimated from half the argument of I_{20}^s , (24).

$$\omega_0 = \sqrt{-\tan\left(\frac{1}{2} \angle I_{20}^s\right)} \quad (28)$$

The magnitude of I_{20}^s reaches its upper bound, I_{11}^s , iff TLS error vanishes. Thus, the magnitudes of (26)-(27) can be utilized as quality measure, q , of model-fitting, e.g. alone $q = |I_{20}^s|$ or in combination, $q = |I_{20}^s|/I_{11}^s \leq 1$ in practice.

We have discretized the derivatives in (26)-(27) by setting τ to σ_{in}^2 and integration domain to $[\omega_{\max}^{-2}, \omega_{\min}^{-2}]$ (14),

$$\frac{d\sigma_{in}^2}{d\tau} = \frac{d\sigma_{in}^2}{d\sigma_{in}^2} = 1 \quad (29)$$

$$\frac{d \log I_{11}}{d\tau} = \frac{1}{\Delta \sigma_{in}^2} \log\left(\frac{I_{11}(\mathbf{r}, \sigma_{in}^2 + \Delta \sigma_{in}^2)}{I_{11}(\mathbf{r}, \sigma_{in}^2)}\right) \quad (30)$$

4. Evidence on data with known ground truth

Evaluating accuracy and isotropy

Fig. 1 shows an example test image where we have synthesized 2 planar waves locally [4]. Our waves were realized via $\cos(C \log r) + \cos(C \cdot \varphi)$ where C is integer. The periods of the two waves are the same ($T_0 = 2\pi/\omega_0$) and change linearly with spatial radius r , attaining $T_0 = 4$ and $T_0 = 17$ at the boundaries of the disc.

On the right half, normally distributed zero mean Gaussian noise is added to the signal, amounting to 65 % of the resulting amplitudes (peak-to-peak) in the average, (0 % in the left-half).

In the figure (top-right), local estimates of the period, T_0 , are shown as a gray image (bright=large T_0), computed from frequency estimations ($T_0 = 2\pi/\omega_0$). The frequency interval in which to search the local frequency was given as $[\frac{2\pi}{13}, \frac{2\pi}{6}]$, representing fingerprint frequencies. Thereby, the method could be stressed to operate under “false” premises, the ground truth span, $[\frac{2\pi}{17}, \frac{2\pi}{4}]$.

The resulting image suggests directional isotropy in the estimations, and the periods increase (brighter) with increased radius. Translated to locations in the image, the apriori interval points in a distance between $[r(T_{\min}), r(T_{\max})]$ from the origin. These are marked as small and medium streaks.

In bottom left, Cr (black) represents the local period estimated along the horizontal diameter (shown in top-right). In the apriori interval of the clean (left) half, the estimation coincides with the ground truth graph, GT, (green) well. In the noisy (right) half, the method offers a good concordance with GT inside the apriori range, changing visibly linearly, although with a slight inclination bias compared to GT. This is most noticeable at points having large T_0 , outside of the apriori interval. The bias is caused by the absolute frequencies of the noise, shifting the estimations systematically. Accordingly, in the noisy part no frequency method should be blamed for having bias but for not having linear change or for not having isotropic behaviour (along circles).

At bottom right, the black graph is the local period T_0 but along the circular demarcation (green, top-right). The ground truth (GT) period is constant and the estimations follow it well in the clean half-circle, and with larger deviations in the noisy part. In the clean half-circle, the mean is a horizontal straight line, which also agrees with GT, (no bias). In the noisy half there is a larger deviation and with a bias (compared to GT), explained by the presence of noise.

The third row of the table beneath represents the relative deviation of the current method, $\hat{\epsilon}_r = |T_{Cr} - T_{GT}|/T_{GT}$, averaged over all half-circles (top-right image), inside and outside of the apriori range, in clean and noisy parts, respectively. The errors reach as low as 0.00 %.

These results support the view that accurate local frequency estimations are achievable by the suggested method

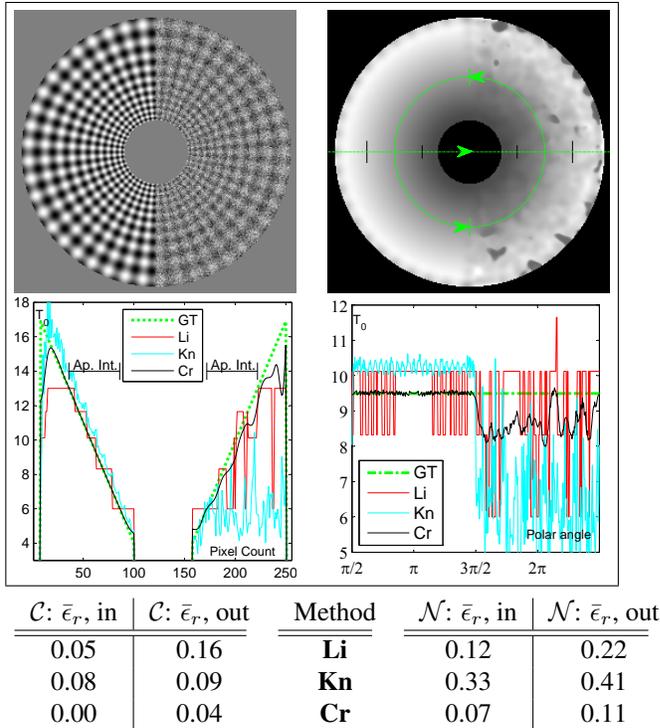


Figure 1. *Top-Left* Two orthogonal waves at each pixel with changing absolute frequencies. *Top Right* Period estimation by the suggested method, Section 3. *Bottom Left* Estimation on mid-horizontal line. *Bottom right* Estimation on a circular path

even without precise knowledge of the frequency range.

Laplacian of Gaussian scale space,

We have implemented the automatic scale selection of [14] to estimate the local frequency. The presumptions (input image, noise, apriori knowledge, scale taps, etc) were identical to those of ours.

The table beneath Fig. 1 (marked as “Li”) suggests that the technique offers a correct estimation of the local period up-to scale discretization step. The frequency estimation in the clean part can be as much as 5 % erroneous compared to the ground truth. We recall that the ground truth frequency is known up to image grid discretization since the original image is synthesized by waves the local frequencies which are known, see C at the beginning of this section. The calculated ground-truth is shown as dashed green curves in the graph. The argmax operation used in the method chooses the closest sampling tap. In the noisy part the frequency estimations are more biased towards the included noise, and with higher variation than the present method.

It can be argued that, our method is in a more advantageous position when the noise type is additive and Gaussian. However, changing the noise to *salt&pepper replacement* noise (not detailed here) continues to support the evidence for excessive vulnerability of the argmax combined with max operation (used in [14]) to noise, when estimating

frequency.

Quadrature filters

In [13] it has been shown that frequency can be estimated by quotients of magnitude response averages of quadrature filters, which we have implemented in the Fourier domain. We report our results emanating from a version of the filters having frequencies tuned to $\{\pi 2^{-k}\}_{k=0}^4$, and orientations tuned to $\{\pi/4\}_{l=0}^3$, Fig. 1. The filter relative bandwidths were $B = 2\sqrt{2}$.

In the two graphs and the table of Fig. 1, we present the results of this technique (marked as Kn). In the left graph, the method is nearly linear in the clean part. However, it suffers from bias in inclination compared to the ground truth. In the noisy part however, the method is not able to retain the linear change, nearly explicitly as Cr and Li methods. The right graph suggests that there is a bias and systematic non-isotropy (periodicity) in the clean part whereas in the noisy part the estimated period is too unstable and far from the linear progression of the ground-truth compared to Cr and Li.

We attempted to alleviate the bias by calibrating the estimated frequencies by a linear as well as affine mapping, which gave promising results on single orientation images (not detailed here) but the calibration constants had to be changed when the setup was exposed to two or more orientation neighborhoods. However, manual determination of the number of orientations upfront is not feasible in many applications.

5. Continuous ridge-counting in fingerprints

In fingerprint identification ridge-counting on paths, e.g lines between minutia points, is a desirable but difficult operation performed by human experts. The difficulty is due to the low-quality of the ridges as well as the tediousness of the task which grows quadratically with the number of points between which the counts are desired.

Our algorithm for ridge-counting is summarized by

$$N_{\mathcal{P}} = \frac{L}{2\pi} \int_{\mathcal{P}} \hat{q}(\tau) \omega(\tau) |\cos(\theta(\tau) - \theta_{\mathcal{P}}(\tau))| d\tau \quad (31)$$

where τ parametrizes the integration path, \mathcal{P} . The function θ is the direction of the local image. Furthermore, $0 < \hat{q}$ is a quality function, which is normalized along the path of integration, (so that $\int \hat{q}(\tau) = 1$). The quality function is assumed to have been generated by local computations/assessments performed along the path. The angle $\theta_{\mathcal{P}}$ is the direction angle of the tangent of the path and L is the length of the path, the presence of which is due to that \hat{q} is normalized, clarified below. Examples of \hat{q} along the path are certainty of frequency, certainty of orientation or a combination of the two.

The equation introduces *continuous ridge-count*, $N_{\mathcal{P}}$, along an arbitrary path. The continuous refers to the fact

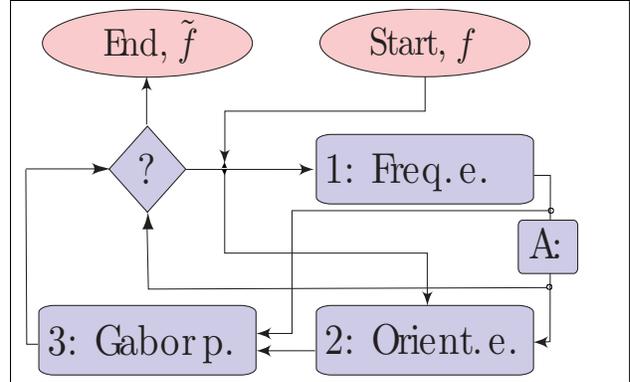


Figure 2. Iterative improvement of dense maps

that $N_{\mathcal{P}}$ is not an integer, i.e. it changes continuously if the path is prolonged continuously, allowing to measure the distance in terms of the average local period. Evidently, $N_{\mathcal{P}}$ can be rounded off to the nearest integer when needed.

The Rationale behind the quality measure in (31) is to improve the accuracy of ridge-counting. Extrapolating ridge-counts for subsets of the path where ridge information is absent is facilitated by \hat{q} .

6. Iterative improvement of dense maps

Estimating dense frequency and orientation maps in poor quality images, including forensic fingerprints, is expected to be challenging. To meet the challenge, we adopted the following iterative improvement scheme consisting of three steps, Fig. 2.

First, frequency estimation was implemented based on scale space of $\log(I_{11})$, as detailed in Section 3.

Second, the estimated frequency map was averaged over to automatically obtain a global inner scale parameter, σ_{in}^2 for the entire image. Using this parameter, a dense orientation estimation by Structure Tensor, I_{20}/I_{11} , was obtained, (3) and (5).

Third, a Gabor filter with frequency and direction parameters estimated as in the previous steps was manufactured for each location in the image. The local image was projected on the real part of the (dynamically computed) Gabor filter².

The projection produces a real image, which is the original image smoothed along its iso-gray curves, therefore here referred to as the enhanced image. At this point, we replace the initial input image with the enhanced image and restart from the first step above. The three steps are iterated until the weighted average of the frequency, A : in the flow-chart, converges, or the maximum number of enhancement cycles (5 here) are reached. Here, the weights were chosen as the

²This estimates the value of a sinusoid at the correct phase w.r.t. the nearest ridge, given that local frequency and orientation are correctly estimated. Derivation details and motivation of this projection are presented in the supplementary material.

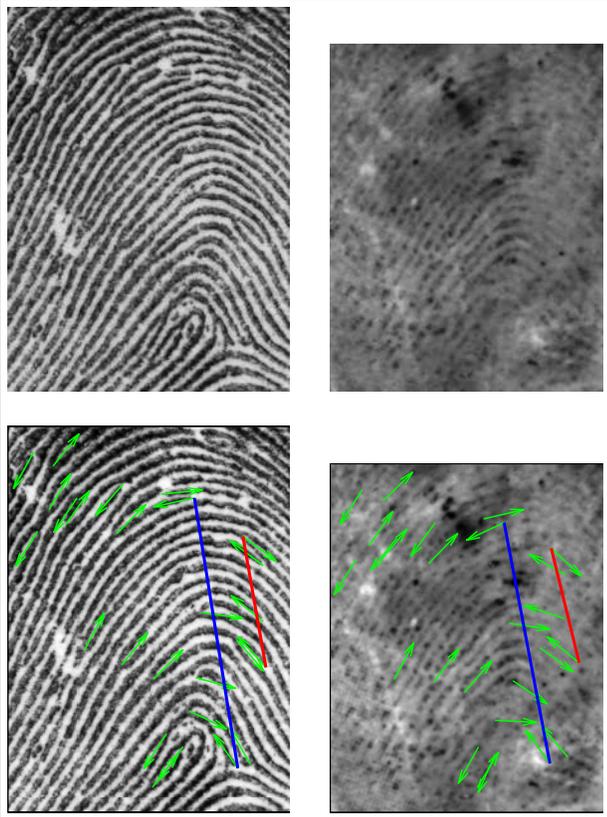


Figure 3. (Top-left) A tenprint detail and (Top-right) its corresponding fingerprint from an image pair in SD27. (Bottom) Same images with common minutia identified by forensic experts, overlaid. Overlaid are also two example minutia edges (red and blue)

certainities of frequencies, but can be other regions of interest, e.g. expert delineations.

7. Evidence on tenprint-fingerprint pairs

We have applied the iterative scheme of Section 6 to images of SD27. In Step 1, we have used the same apriori knowledge $\omega_0 \in [\frac{2\pi}{13}, \frac{2\pi}{6}]$, but using 3 discrete scales, $\sigma_{in}^2 : 0.92, 2.60, 4.28$. For the majority of images of SD27, at most 3 iterations were sufficient for convergence. This is exemplified by Fig. 4, for the converged frequency and orientation maps, and by Fig. 5 for converged enhanced image. In the applications below, only the dense maps were used whereas the enhanced images were not used explicitly.

Ridge counting experiments

We have applied our ridge counting method, Section 5, on paths of tenprint-fingermarks of SD27. As paths, we have chosen the lines joining pairs of minutia, called here *edges*, e.g. the blue and red lines in Fig. 3.

Because we have 74278 mated edge-pairs, whose end minutia were verified by experts (Match-Set), and that the respective ridge counts on them must agree, we can evalu-

ate our automatic ridge counting by computing to what extent the ridge counts on mated edges of tenprint-fingerprint pairs agree, a correlation. The rationale behind this evaluation is that if all edge-pairs would have the same (Euclidean) Lengths, then matching two edges by chance is $1/74278$.

Table 1 depicts the relative agreements $R_A(|\epsilon_{rc}|)$ between our automatic estimations of ridge counts on corresponding edges, rounded off to closest integers. For 78 % of the edges (57937), the agreement between the ridge counts were at most one ridge apart, when our frequency and orientation maps in combination with our ridge counting method was in use (Row of R_{Cr}). The columns depict the relative agreements measured at different error tolerances (0 pixels, 1 pixels, etc.).

We have used all edges exhaustively and blindly. However a human expert avoids some edges since they would be more prone to counting errors than others *constellation errors*, described next. The first type of this error occurs if an edge between two minutia passes through or nearby a third minutia. The total ridge count on the edge can then be one more or less, depending on the localization accuracy of the (three) involved minutiae. If such an edge is barred from ridge counting, the (smaller two) edges joining the middle minutia to the initial two still contain the same information for identification purposes.

In Fig. 3, drawn in red, an edge suffering from this type of error is shown, along with another edge, shown in blue which does not suffer from the phenomenon. On red edges, the (automatic) ridge counts differ with 2.7 ridges (7.9, 5.2 for tenprint, fingerprint respectively), whereas along the blue edges, the same counts differ with 0.0 ridges (16.6, 16.6). Even if the (blue) edge is long and runs through noisy ridges, the suggested method was able to estimate the ridge count better than on the shorter (red) edge.

In Table 1, marked as R'_{Cr} , agreements between the ridge counts are shown the underlying minutia edges had been screened. The latter was done by rejecting edges with a third minutia within 8 pixels (Euclidean) distance, resulting in 6045 edges. The 8 pixels threshold corresponds roughly to the average inter-ridge distance. Accordingly, in 83% of mated edges, the ridge counts were in agreement. On the average this represents 26 different ridge counts per tenprint-fingerprint pair which can be offered to experts automatically in SD27.

In our ridge counting, the quality measures of dense maps offer interpolation in intervals where ridges are not available from adjacent intervals where they are available. We are not aware of other studies on ridge counting for fingerprints to do comparison. Instead, we have used a different frequency map, Section 4, but with our orientation map and ridge counting to evaluate the role of frequency maps. The results, marked with Li in Table 1, support the view that our frequency maps to ridge counting is preferable.

Table 1. Prevalence of $|\epsilon_{rc}|$, on edges between minutia (SD27)

$ \epsilon_{rc} $	0	1	2	3	4	5
$R_{Li}(\epsilon_{rc})$	0.37	0.71	0.84	0.91	0.94	0.96
$R_{Cr}(\epsilon_{rc})$	0.43	0.78	0.90	0.95	0.97	0.98
$R'_{Li}(\epsilon_{rc})$	0.43	0.77	0.88	0.93	0.94	0.96
$R'_{Cr}(\epsilon_{rc})$	0.51	0.83	0.92	0.96	0.97	0.98

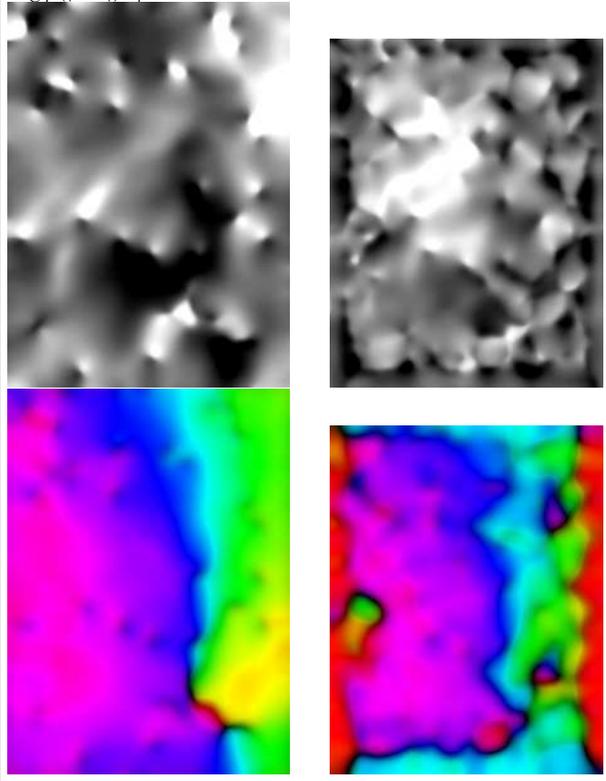


Figure 4. (Top) Dense frequency maps of the tenprint-fingermark pair in Fig. 3. Bottom Dense orientation map of the same pair. Both maps constitute the (only) inputs of dense ridge counting algorithm along arbitrary paths

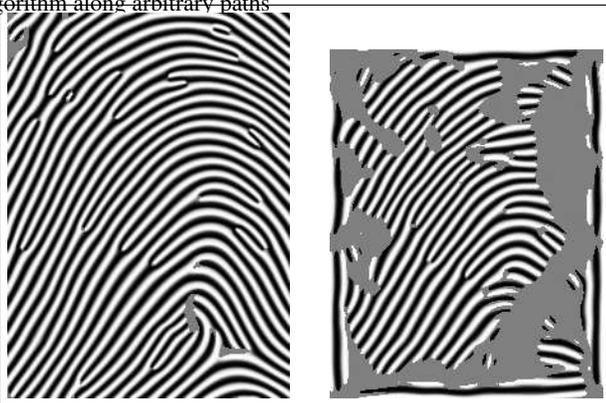


Figure 5. Enhanced images (of Fig. 3) using dense orientation and frequency maps. Gray regions are formed by points where frequency and orientation certainties were too low

Frequency map as descriptor of minutia

Our purpose is to evaluate the description power of frequency maps around single minutiae in a fingermark. This

is measured by verifying the frequency map around a minutia in a fingermark against its mate in the genuine tenprint, as well as against its non-mates selected from a large random set of minutiae neighborhoods from tenprints.

In Fig. 6 a fingermark (SD27) left on a cheque along with its 106 minutia identified by experts are shown (left) along with the zoomed-in minutia area (right). The periodical strokes and letters in high curvature ridge areas present additional deviations from a single frequency and/or single orientation model.

The match score s_M used in the verification has been obtained by $s_M = \langle f, g \rangle / (\|f\| \|g\|)$ where f , and g where $s_M \leq 1$ are the frequency maps around the two minutiae, after (rotational) alignment. A rotation is applied to the tenprint neighborhood of the minutia such that the directions of the two minutiae agree. In SD27 1032 fingermark minutiae labels are available with mates in matching tenprints [16] so that as-many client verifications can be made. Consequently, even non-mate minutiae are available, choosable at random. We have effectuated random impostor verifications against all genuines, 5508 total. We emphasize that both client and impostor refer to the identity of a single minutia (not the finger).

In the experiments, f and g were chosen to be maps extracted from neighborhoods of various sizes and shapes. We only report on two sizes and shapes that provide most insight. First is a *disc* neighborhood consisting of (digital grid) points with distance less than 23 pixels to the origin (minutia location). Second is a *ring* consisting of points with distance between 67 and 90 pixels to the origin.

Figure 6 shows the client (false rejection, FR) versus impostor (false acceptance, FA) probability estimates, drawn as Detection Error Trade-off (DET) curves. These are equivalents of Receiver Operator Curves but with “logarithmed” axes. The graphs on the disc (DS-PR), on the ring (RN-PR), and as the fusion of both (FS-PR) are marked, yielding the Equal Error Rates (EER) of 40 %, 32% and 30% respectively.

We have also used frequency maps produced by the method of [2]. The corresponding graphs are marked as DS-CH, RN-CH, and FS-CH, yielding EERs of 48 %, 39% and 39% respectively. The results suggest that 9 (=39-30) better decisions were made in 100 trials when our frequency maps were in use.

The model of [2] fits a (single) frequency (ring) to the power spectrum for a whole block. Accordingly, the method represents a lower resolution of the frequency map compared to ours where the frequency is the expectation of the radial coordinate when the power spectrum is interpreted as probability. Our frequency estimation is a TLS fit in the log scale space, and affords a higher spatial resolution.

Heuristics, for example continuity of orientation, and frequency can be added to any frequency estimation method

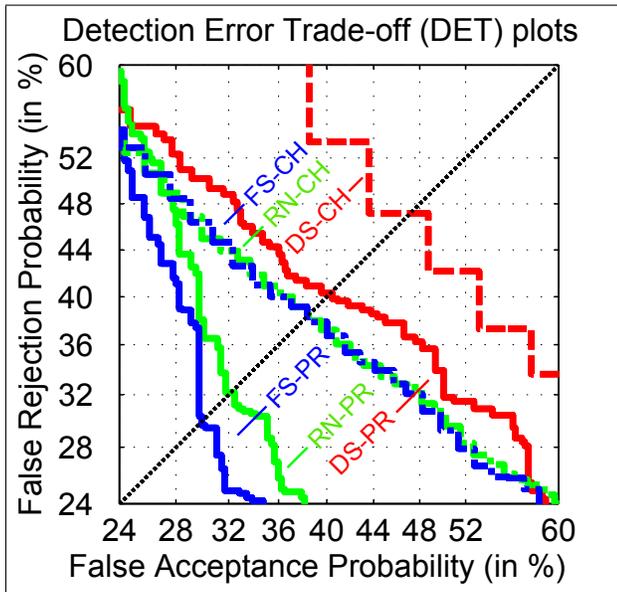


Figure 6. (Top) A sample fingermark with contamination superimposed with its minutiae and its zoomed version. (Bottom) Impostor versus Client error probabilities for minutia identities based on frequency maps estimated on SD27.

(i.e. to both methods) to reduce the problems caused by ambiguities when a foreign wave pattern contaminates a genuine wave, e.g. as is done in palm-print recognition [10] (their eq. (5) and (6)), or by using robust statistics to reject outliers. Additionally, fingerprint minutia information can be combined automatically or manually to deliver more robust identification of a fingermark in real crime cases.

8. Conclusions

Our study suggests that more accurate and isotropic frequency estimations are possible by using a novel non-linear scale space compared to state of the art.

The study suggests that frequency measurement is an orientation estimation problem in the logarithmic scale space generated by the Structure Tensor.

We have presented a new method counting ridges continuously in tenprints as well as fingermarks based on dense frequency and orientation maps. Experimental evidence supports that it can count ridges without edge thinning, ridge

extraction, etc, at agreement level of 83% between mated paths.

Experimental evidence supports that our frequency map is more helpful to verify mated minutiae neighborhoods than published frequency map methods used in fingerprint recognition.

The presented evidence supports that with 3 to 5 sampling taps in the scale-space, one can estimate the frequency accurately and obtain a certainty of the estimation.

Independently obtained dense frequency and orientation maps can be combined in iterative schemes to improve each other.

The theory supported by 2 theorems suggest that these benefits can be extended to images with higher dimensions than 2 for other applications.

References

- [1] J. Bigun and G. Granlund. Optimal orientation detection of linear symmetry. In *ICCV, London, June 8–11*, pages 433–438. IEEE Computer Society, 1987.
- [2] S. Chikkerur, V. Govindaraju, and A. N. Cartwright. Fingerprint image enhancement using stft analysis. In *Pattern Recognition and Image Analysis*, pages 20–29. Springer, 2005.
- [3] S. Di Zenzo. A note on the gradient of a multi-image. *Computer Vision, Graphics, and Image Processing*, 33(1):116–125, 1986.
- [4] M. Felsberg and G. Sommer. Image features based on a new approach to 2D rotation invariant quadrature filters. In A. H. et. al., editor, *Computer Vision—ECCV 2002*, LNCS 2350, pages 369–383. Springer, Heidelberg, 2002.
- [5] W. Förstner and E. Gülch. A fast operator for detection and precise location of distinct points, corners and centres of circular features. In *Proc. Intercommission Conference on Fast Processing of Photogrammetric Data, Interlaken*, pages 281–305, 1987.
- [6] L. Hong, Y. Wand, and A. Jain. Fingerprint image enhancement: Algorithm and performance evaluation. *IEEE-PAMI*, 20(8):777–789, 1998.
- [7] P. Hough. *Method and means for recognizing complex patterns*. U.S. patent 3,069,654, 1962.
- [8] D. Hubel and T. Wiesel. Receptive fields of single neurons in the cat’s striate cortex. *J. physiol. (London)*, 148:574–591, 1959.
- [9] A. Jain and J. Feng. Latent fingerprint matching. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 33(1):88–100, 2011.
- [10] A. K. Jain and J. Feng. Latent palmprint matching. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 31(6):1032–1047, 2009.
- [11] M. Kass and A. Witkin. Analyzing oriented patterns. *Computer Vision, Graphics, and Image Processing*, 37:362–385, 1987.
- [12] H. Knutsson. Representing local structure using tensors. In *Proceedings 6th Scandinavian Conf. on Image Analysis, Oulu, June*, pages 244–251, 1989.

- [13] H. Knutsson, C.-F. Westin, and G. Granlund. Local multi-scale frequency and bandwidth estimation. In *Image Processing, 1994. Proceedings. ICIP-94., IEEE International Conference*, volume 1, pages 36–40. IEEE, 1994.
- [14] T. Lindeberg. Feature detection with automatic scale selection. *International Journal of Computer Vision*, 30(2):79–116, 1998.
- [15] D. Maio and D. Maltoni. Ridge-line density estimation in digital images. In *International Conference on Pattern Recognition*, pages Vol I: 534–538, 1998.
- [16] A. Mikaelyan and J. Bigun. Ground truth and evaluation for latent fingerprint matching. In *CVPR Workshop on Biometrics*. IEEE-XPLORE, Piscataway, NJ, June 18-21, 2012.
- [17] K. Nilsson and J. Bigun. Prominent symmetry points as landmarks in fingerprint images for alignment. In *ICPR-16 International Conference on Pattern Recognition*, volume III, pages 395–398. IEEE computer society, Aug. 11-15, Quebec, Canada 2002.
- [18] A. Witkin. Scale-space filtering. In *8th Int. Joint Conf. on Artificial Intelligence, Karlsruhe, Aug. 8–12*, pages 1019–1022, 1983.