A Rotation-extended Cepstrum Technique and Its Application to Medical Images

T. M. Lehmann, C. Goerke*, A. Kaupp, W. Schmitt***, and R. Repges***

Institute of Medical Informatics and Biometry Medical Faculty, Aachen University of Technology (RWTH), Aachen, D-52057 Germany
Phone: +49 / 241 / 80 - 88793
Fax: +49/241 /8888 - 426
E-mail: lehmann@vaire.imib.rwth-aachen.de
Url: http://www.imib.rwth-aachen.de/www/mitarb/lemmi
** Institute of Medical Informatics and Biometry, Aachen University of Technology (RWTH), Aachen, D-52057 Germany*
*** Institute for Measurement Technology, Aachen University of Technology (RWTH), Aachen, D-52056 Germany*
**** Clinic for Oral, Maxillofacial, and Facial Plastic Surgery, Aachen University of Technology (RWTH), Aachen, D-52057 Germany*

Abstract—Spatial registration is a major problem arising whenever several images of similar contents are to be compared. Considering translations only, two-dimensional cepstral techniques have been proven to be exact and robust against noise or intensity variations. Furthermore, cepstral filtering is numerically more efficient than most common approaches to image registration based on cross-correlation or template matching. In a previous paper [1], we proposed a cepstrum-based two-dimensional matching technique accessing rotations and translations. The logarithmic polar transform of the power spectra of both images to be registered is used for the decoupling of rotations and translations (similar to the Fourier–Mellin transform). Rotations are detected first matching the mapped spectra by two-dimensional cepstrum analysis. After rotating back one image, the relative shift is determined using the same cepstrum technique. In clinical practice, the rotation detection step was discovered as the weakness of this registration technique. Based on 855 pairs of dental radiographs acquired in *a priori* known positions, three different approaches of matching the mapped spectra are compared: the cepstrum technique, the cross-correlation, and the entropy of the one-dimensional histogram distribution function of the subtraction image. The combination of the log–polar mapped power spectra of both X-rays with the entropy-measure allows the best detection of rotations. The union with common cepstrum methods correcting translations results in a robust rotation-extended cepstrum technique. The usefulness of this new technique is demonstrated on several applications in medical imaging. Dental radiographs, acquired intraorally from the same patient and the same dental region but in large time distances, are registered and subtracted allowing densitometric bone assessment. Video sequences showing the vibration of human vocal cords are corrected for distortions caused by the movement of patient or examiner. Histological slices of human ankles are registered before 3D reconstruction. In all these examples the contents of images to be registered differ locally.

1. INTRODUCTION

Spatial registration of multispectral or multitemporal digital images is a major problem arising whenever data of a recording system have to be compared, and is therefore a focus of recent discussion in image processing [2, 3]. To obtain a comparison on a pixel by pixel base, it is necessary to spatially register the images and correct the geometric distortions, e.g., translational and rotational movements, with sufficient accuracy. A matching algorithm should be robust with respect to noise and accurate even if the same pattern has different intensities and/or different positions and orientations in the two images to be registered. Furthermore, a matching algorithm should be numerically efficient.

The most common matching approach is based on the cross-correlation function and is also referred to as template matching [4]. It is derived assuming white noise distorting the object in the image. The size of the object determines the template's dimension. It corresponds directly to the computing time required for the calculation of all possible values of the parameters describing the geometric transform relating the object to its representation in the reference image. While such an extensive calculation is extremely time consuming, moments or moment invariants based matching techniques are used [5], especially if further distortions than translations can be expected. Those techniques are often applied to the binary representation of the object (e.g., character recognition) or the object's outlines.

In the field of medical image processing, the object/background model is invalid, i.e., the entire

Received January 20, 1996

Pattern Recognition and Image Analysis, Vol. 6, No. 3, 1996, pp. 592–604.

image is the object. In this case, several alternatives to the cross-correlation function have been proposed. In one group of matching techniques, the cross-correlation is exchanged by other correlation-like functions, e.g., the standard deviation of the difference image [6, 7], or the number of sign changes along the lines of the difference image [8, 9]. Since all these techniques at least require the geometric transform of one image for all possible values of the parameters, those algorithms are numerically inefficient.

Another group of matching techniques bases on Fourier spectral or cepstral analysis. The two-dimensional cepstral techniques have been proven to be very accurate and characterized by an outstanding robustness against uncorrelated noise and intensity distortions [10]. In addition, cepstral filtering requires no segmentation of the images. However, the classical way of cepstrum analysis is restricted to the detection of translational shifts only and the pictures are not allowed to differ in rotation. Therefore, an extension to the detection of both translation and rotation was proposed by the authors in a previous paper [1].

In this paper, we describe the optimization of the rotation-extended cepstrum technique and its application in medical imaging. Since high-resolution CCD sensors are available for intraoral X-ray imaging [11, 12], our *in vitro* study bases on dental radiology. To optimize the rotation detection step, a total of 855 different pairs of radiographs with *a priori* known misalignments were systematically analyzed. We compared the results of three approaches: the cepstrum filter, the cross-correlation, and the entropy of the one-dimensional histogram function of the subtraction of the mapped spectra.

The next section introduces the Fourier and the Mellin transforms and their combination to an invariant image descriptor. All three techniques use the decomposition of rotation and translation by this descriptor. Section 3 reviews the cepstrum technique for detection of translational shifts and the three extensions to rotations. Computational aspects and efficient implementation using the Hartley transform are described in Section 4. Our method acquiring X-rays with *a priori* known geometric positions and the results registering 855 pairs of dental radiographs are succeedingly documented in Section 5. The application of the optimized technique to medical X-rays, videoscopic, and histological images and sequences is described in Section 6.

2. FOURIER–MELLIN INVARIANT IMAGE DESCRIPTOR

In this section, the Fourier and Mellin transforms are introduced and combined to a rotation-, scaling-, and translation- (RST-) invariant image descriptor. It is shown that this descriptor can be used to decouple RST movements.

2.1. Fourier Transform

The integral transforms [13]

$$F(\omega) = \int_{-\infty}^{\infty} f(t)e^{-j\omega t}dt \text{ and}$$
$$f(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} F(\omega)e^{j\omega t}d\omega \quad (1)$$

are well known as the Fourier and inverse Fourier transform, respectively. One can easily show that the magnitude of the two-dimensional Fourier transform $|F(u, v)|$ of the image function $f(x, y)$ is invariant to translations in the image plane, while rotations in the (x, y) plane are transformed to equal rotations in the (u, v) plane, and scaling is inverted [13], e.g., expansions are transformed to shrinkings.

2.2. Mellin Transform

The Mellin transform $M(s)$ of a continuous one-dimensional function $f(z)$ is given by [14, 15]

$$M(s) = \int_{0}^{\infty} f(z)z^{s-1}dz \text{ and}$$
$$f(z) = \frac{1}{j2\pi}\int_{c-j\infty}^{c+j\infty} M(s)z^{-s}ds \quad (2)$$

denoting the inverse Mellin transform, respectively, where c is the real part of the complex variable $s = c + j\omega$. The Mellin transform is usually used to compute the moments of $f(z)$ [13]. For example, area, first, and second moments of $f(z)$ are equal to $M(1)$, $M(2)$, and $M(3)$, respectively, and the second moment about centroid is $M(3) - [M(2)]^2/M(1)$.

Substituting z by e^{-t} we obtain $\ln(z) = -t$ and $dz/dt = -e^{-t}$. Furthermore, the integration limits in (2) transforms like $z \longrightarrow 0 \Rightarrow t \longrightarrow \infty$ and $z \longrightarrow \infty \Rightarrow t \longrightarrow -\infty$ and the Mellin transform becomes

$$M(s) = \int_{\infty}^{-\infty} f(e^{-t})e^{-t(s-1)}(-e^{-t})dt$$
$$= \int_{-\infty}^{\infty} f(e^{-t})e^{-st}dt = L(s), \quad (3)$$

with $L(s)$ denoting the two-sided Laplace transform of a function with distorted coordinates $f'(t) = f(e^{-t})$. Further substituting $s = j\omega$ in (3) yields (1). Therefore, the undamped Mellin transform

$$M(\omega) = \int_{0}^{\infty} f(z)z^{-j\omega-1}dz = \int_{-\infty}^{\infty} f(e^{-t})e^{-j\omega t}dt \quad (4)$$

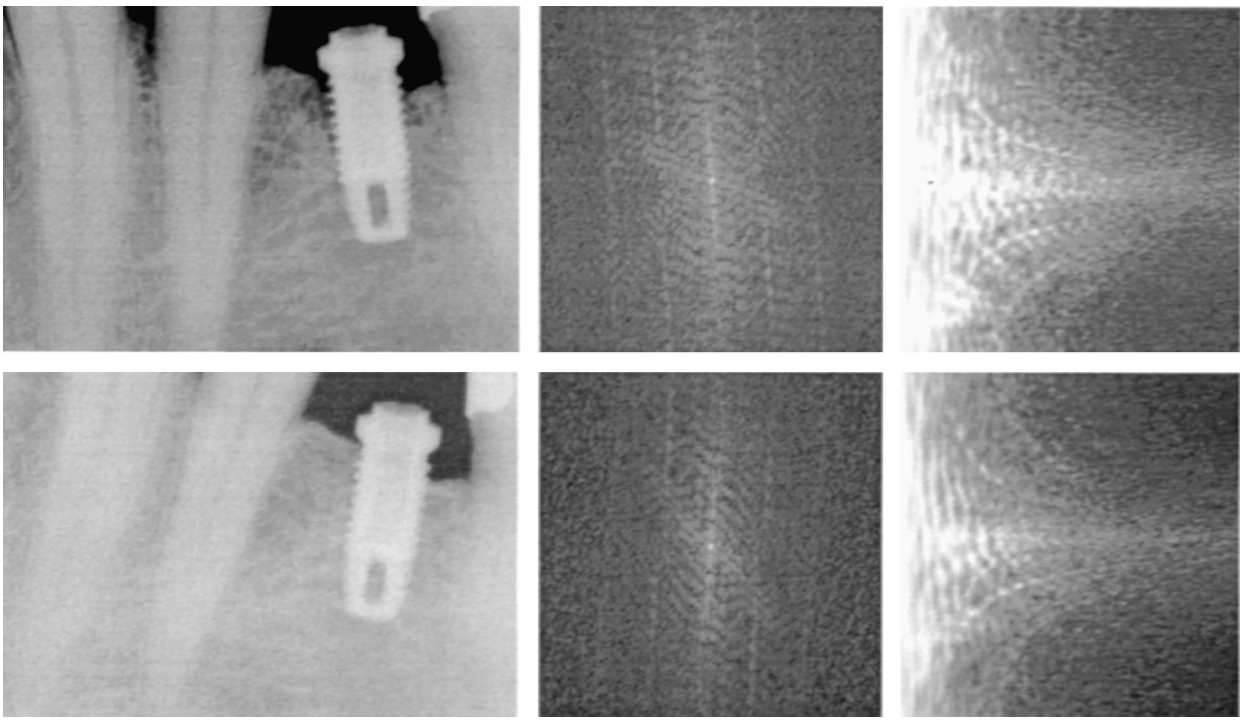

Fig. 1. Logarithmic polar mapping of Fourier power spectra. The images $f(x, y)$ on the left show dental X-rays. Teeth as well as implants are figured. The Fourier power spectra $|F(u, v)|$ are shown in the middle. Note their invariance to translations while their rotational orientation corresponds to that in the spatial domain. The results of logarithmic polar mapping $|F(\rho, \varphi)|$ are displayed on the right in each line, respectively. Rotations are transformed into shifts along the vertical axis.

is computable by the Fourier transform (1), if the coordinate $t \longrightarrow \ln(1/t)$ is logarithmically deformed at first,

$$t \equiv e^{-\left(\ln\left(\frac{1}{t}\right)\right)} \equiv e^{-(-\ln(t))}. \tag{5}$$

Let f_1 and f_2 be two functions differing only in scale $f_2(z) = f_1(\alpha z)$. Replacing αz by τ, $z = \tau/\alpha$, $dz = d\tau/\alpha$ in the Mellin transform (4) we obtain

$$\begin{aligned} M_2(\omega) &= \int_0^\infty f_1(\alpha z) z^{-j\omega-1} dz \\ &= \int_0^\infty f_1(\tau)\tau^{-j\omega-1}\left(\frac{1}{\alpha}\right)^{-j\omega-1}\frac{1}{\alpha}d\tau = \alpha^{j\omega}M_1(\omega). \end{aligned} \tag{6}$$

One can easily recognize that because of $|\alpha^{j\omega}| = \left|e^{j\omega\ln\alpha}\right| = 1$ equation (6) holds $|M_1(\omega)| = |M_2(\omega)|$. Therefore, the magnitude of a Mellin transformed function is invariant to scales.

2.3. RST-Invariant Image Descriptors

An overview of invariant image descriptors is given elsewhere [5]. In the following, the Fourier and the Mellin transformations are combined to an RST-invariant image descriptor similar to that introduced by Schalkoff [16] or Haken [17]. Therefore, let f_1 and f_2 be two given functions with f_2 being a (x_0, y_0) shifted, α rotated, and β scaled version of f_1,

$$\begin{aligned} f_2(x, y) = f_1(&\beta(x\cos\alpha + y\sin\alpha) - x_0, \\ &\beta(-x\sin\alpha + y\cos\alpha) - y_0). \end{aligned} \tag{7}$$

The Fourier power spectra are computed to

$$\begin{aligned} &|F_2(u, v)|^2 \\ &= \left|\frac{1}{\beta^2}F_1\left(\frac{u\cos\alpha + v\sin\alpha}{\beta}, \frac{-u\sin\alpha + v\cos\alpha}{\beta}\right)\right|^2. \end{aligned} \tag{8}$$

This representation of the functions f depends only on the angle of rotation α and the amount of scale β. Both may be decoupled applying a transformation to polar coordinates (Fig. 1). Taking into account the addition theorems of trigonometric functions, the substitution $u = r\cos\varphi$ and $v = r\sin\varphi$ in (8) yields

$$\begin{aligned} &|F_2(r\cos\varphi, r\sin\varphi)|^2 \\ &= \left|\frac{1}{\beta^2}F_1\left(\frac{r\cos(\varphi - \alpha)}{\beta}, \frac{r\sin(\varphi - \alpha)}{\beta}\right)\right|^2. \end{aligned} \tag{9}$$

Denoting the power spectra in (9) as functions of (r, φ) leads to

$$|F_2(r, \varphi)|^2 = \left|\frac{1}{\beta^2}F_1\left(\frac{r}{\beta}, \varphi - \alpha\right)\right|^2. \quad (10)$$

Scalings are transformed into stretches of the r-axis, while rotations are mapped to shifts along the φ-axis. Applying the scale-invariant Mellin transform (4) to the r-coordinate and the shift-invariant Fourier transform (1) to the φ-coordinate (Fourier–Mellin transform) equation (10) results in an RST-invariant image representation of both functions f_1 and f_2. As shown in Section 2.2, the Mellin transform can be calculated using the Fourier transform if the coordinate $\rho = -\ln(r)$ is mapped logarithmically. Combining (5) and (10) yields

$$|F_2(\rho, \varphi)|^2 = \left|\frac{1}{\beta^2}F_1(\rho + \ln(\beta), \varphi - \alpha)\right|^2. \quad (11)$$

In this representation of the entire images f_1 and f_2 the RST movements are decoupled. Translational shifts are eliminated while magnifications are represented as shifts in the p-axis and rotations are expressed as shifts in the φ-axis. A further calculation of the Fourier power spectra of the functions given in (11) results in an RST-invariant image descriptor

$$|M_2(u, v)|^2 \sim |M_1(u, v)|^2. \quad (12)$$

The mapping of multiplications (coordinate-scale) to additions (coordinate-shift) is a well-known feature of the logarithm. Therefore, related methods are often used without being referred to as the Fourier–Mellin transform [1, 2, 18].

3. THE CEPSTRUM TECHNIQUE

The next paragraph recapitulates the classical cepstrum technique for the detection of pure translational shifts. Based on the Fourier–Mellin RST-invariant image descriptors, different methods for the extension to rotational shifts are proposed in the subsequent paragraphs.

3.1. Classical Cepstrum

The cepstrum analysis was developed as a one-dimensional technique to examine seismographic data containing echos of some arbitrary wave packets [19, 20]. In 1975 the (power-) cepstrum was derived again from the homomorphic system theory [21]. Based on the separability of the discrete Fourier transform (DFT) $\mathcal{F}\{\cdot\}$ cepstral filtering can easily be transferred to two-dimensional image processing. In general, the power cepstrum of a function f is the power spectrum of the logarithm of the function's power spectrum. Therefore, the two-dimensional power cepstrum $\mathcal{C}\{\cdot\}$ may be defined in the spatial domain (x, y) as

$$\mathcal{C}\{f(x, y)\} = \left|\mathcal{F}\{\log|\mathcal{F}\{f(x, y)\}|^2\}\right|^2. \quad (13)$$

Using the most common way of cepstral filtering to estimate a scene's translational shift, first of all, the two images to compare with are placed next to next in the center of a so-called cepstrum window. Assume a cepstrum window consisting of the reference image $r(x, y)$ on the left hand side and the subsequent test image $t(x, y) = a_0 r(x - x_0, y - y_0)$ on the right-hand side. Then, the resulting cepstrum window $g(x, y)$ may be expressed like

$$\begin{aligned} g(x, y) &= r(x, y) + t(x - D, y) \\ &= r(x, y)[\delta(x, y) + a_0\delta(x - (D + x_0), y - y_0)], \end{aligned} \quad (14)$$

where a_0 is an amplitude scale factor and D describes the dimension of the square images r and t. The relative translational shifts between both images are written in the spatial x direction in terms of x_0 and in the y-direction in terms of y_0. A stepwise calculation of the power cepstrum (13) shows the logarithm of the power spectrum of the composite signal g containing cosinusoidal ripples whose amplitude and frequency are related to the intensity factor a_0 and the translational displacements $D + x_0$ and y_0

$$G(u, v) = R(u, v)(1 + a_0 e^{-i2\pi(u(D + x_0) + vy_0)}) \quad (15)$$

$$\begin{aligned} \log|G(u, v)|^2 &= \log|R(u, v)|^2 \\ &+ \log(1 + 2a_0\cos(2\pi(u(D + x_0) + vy_0)) + a_0^2) \\ &= \log|R(u, v)|^2 + \log|1 + a_0^2| \\ &+ \frac{2a_0}{1 + a_0^2}\cos(2\pi(u(D + x_0) + vy_0)) \mp \dots, \end{aligned} \quad (16)$$

where $R(u, v)$ and $G(u, v)$ are the Fourier transforms in the frequency domain (u, v) of the reference image and the cepstrum window, respectively. In (16) the power series expansion $\log(1 + x) = \sum_{n=1}^{\infty}(-1)^{n+1}x^n/n$ with $-1 < x \le 1$ is inserted. A second power-spectrum operation is performed to obtain the power cepstrum of $g(x, y)$,

$$\begin{aligned} \mathcal{C}\{g(x, y)\} &= \mathcal{C}\{r(x, y)\} + A\delta(x, y) \\ &+ B\delta(x \pm (x_0 + D), y \pm y_0) \\ &+ C\delta(x \pm 2(x_0 + D), y \pm 2y_0) + \dots, \end{aligned} \quad (17)$$

which equals the power cepstrum of the reference image plus a train of delta functions occurring at integer multiplies of the translational shifts $(D + x_0, y_0)$. Thus, the translational difference between the two images r and t can simply be detected by inspecting the distance between the origin and the location of the first maximum peak in the cepstral plane.

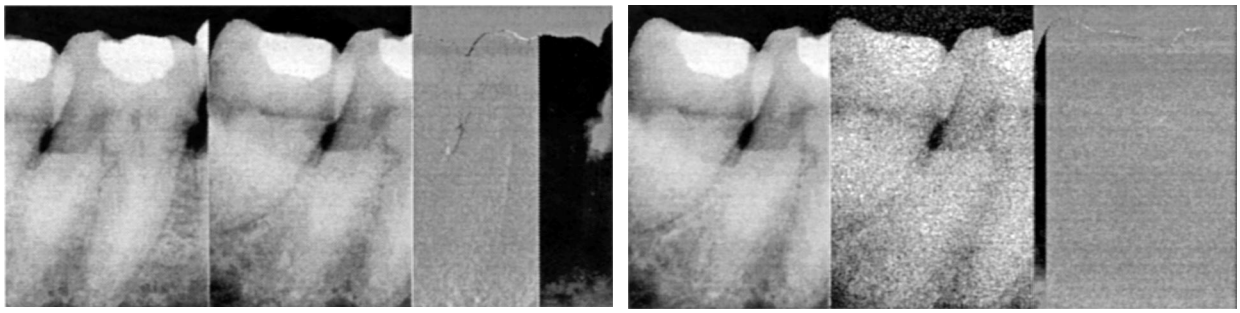

Fig. 2. Registration of translations using the classical cepstrum technique. The figure's left part illustrates the classical cepstrum's property registering very large shifts (up to 50% of the width of images). The original-, the reference-, and the subtraction-image after registration are displayed adjacently from left to right. The dark bar in the difference image results from subtracting non-overlapping regions. The example on the right demonstrates the robustness of the technique with respect to noise (down to signal-to-noise ratios below 1.0) [10].

Although this classical cepstrum technique requires no segmentation and is reliable, efficient, and immune to noise (Fig. 2), it has the disadvantage of being sensitive to rotational changes between the images being registered.

3.2. Rotation-Extended Cepstrum Technique

An extension to rotations of classical cepstrum filtering can be given using the Fourier–Mellin RST-invariant image descriptor. The logarithmic polar mapping of the Fourier power spectra (11) yields an image representation where translations are eliminated and rotations and scales are mapped into shifts. Those may be detected applying the classical cepstrum (17). After appropriately rotating and scaling one of the images, [22, 23] the cepstrum technique (17) may be applied again with $\alpha = 0$ and $\beta = 1$ in (7) now [1].

3.3. Optimization for Medical Imaging

In medical imaging, scales are often neglectable. For example, in dental radiology, the X-ray tube is touching the patient's face and the intraoral film is pressed onto jaw and teeth. The approximation $\beta \approx 1$ in (11) leads to

$$|F_2(\rho, \varphi)|^2 = |F_1(\rho, \varphi - \alpha)|^2, \tag{18}$$

and therefore, the registration step for rotation detection has to be performed only in one dimension. Since cepstral techniques always assess two dimensions, correlation based methods can be easily designed to register only one dimension.

The cross-correlation function

$$\mathrm{CCF}(t) = \sum_{\rho}\sum_{\varphi}(|F_1(\rho, \varphi + t)|^2 |F_2(\rho, \varphi)|^2) \tag{19}$$

can be used efficiently, where the maximum $\mathrm{CCF}_{max} = \mathrm{CCF}(\alpha)$ indicates the rotation angle α to be detected. Nevertheless, a lot of multiplications are required to compute the cross-correlation (19).

The amount of computation is drastically reduced if the similarity measure is constructed from an one-dimensional function. Let $d_t(\rho, \varphi)$ be the subtraction of the functions F_1 and F_2 defined in (18) which can be calculated without any multiplications,

$$d_t(\rho, \varphi) = |F_1(\rho, \varphi + t)|^2 - |F_2(\rho, \varphi)|^2, \tag{20}$$

and t again denoting the misalignment in the φ-axis. Then, the histogram distribution $h_t(p)$ provides a one-dimensional function describing the similarity between F_1 and F_2 depending on the parameter t, where p is in the range of values of d_t. The standard deviation [24] or the entropy [25, 26] of the histogram of the difference image

$$\mathrm{EHDI}(t) = -1\sum_{g} h_t(p)\log(h_t(p)) \tag{21}$$

may be used to evaluate the best adjustment with $\mathrm{EHDI}_{max} = \mathrm{EHDI}(\alpha)$ indicating the rotational angle α to be detected.

4. IMPLEMENTATION

Some preprocessing steps are used to improve the results of cepstrum filtering. These methods are described next while the computational improvement applying the discrete Hartley transform is shown in the last paragraph of this section.

4.1. Preprocessing

The discrete Fourier transform (DFT) assumes the images to be periodic. To avoid artefacts resulting from the nonperiodic property of finite images windowing is required [20]. In our application, the Kaiser–Bessel

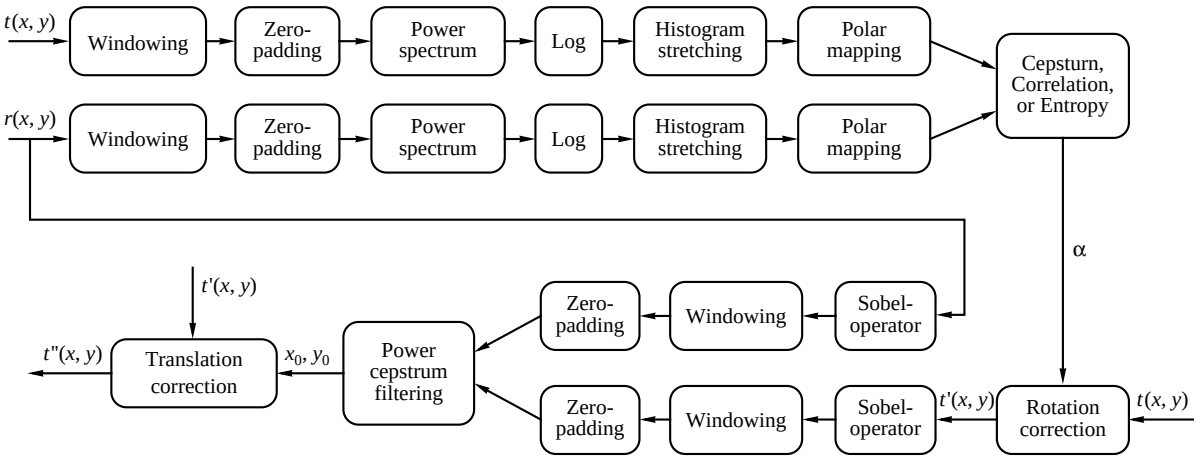

Fig. 3. Rotation-extended cepstrum technique. After windowing and zero-padding, the Fourier power spectra of both images to be registered are calculated. Before mapping to polar coordinates, the spectra are histogram optimized. Rotation is detected first using either the classical cepstrum technique, or the correlation, or the entropy based similarity measure. These are the three methods to be compared. After rerotating one image, the cepstrum technique is applied in a second step also registering translations.

window $w(n)$ for the dimension of the image N, $0 \le n < N$ is approximated by [27]

$$w(n) = b_0 - b_1\cos\left(2\pi\frac{n}{N}\right) - b_2\cos\left(2\pi\frac{2n}{N}\right) - b_3\cos\left(2\pi\frac{3n}{N}\right) \quad \text{with} \quad \begin{cases} b_0 = 0.40243 \\ b_1 = 0.49804 \\ b_2 = 0.09831 \\ b_3 = 0.00122 \end{cases} \tag{22}$$

The algorithms for the fast Fourier transform (FFT) are usually limited to images with dimensions of powers of two. Zero-padding all images before transforming allows the handling of arbitrary dimensions of the image's column or row length. Additional preprocessing by Sobel-filtering is also applied to emphasize edge information [10]. In addition, the Fourier spectra are histogram-optimized by stretching their logarithmic scale before they are mapped into polar coordinates. The resulting block diagram is presented in (Fig. 3).

The logarithmic polar mapping is used to turn rotations into translations along the φ-axis. In order to calculate the logarithmic polar mapping on discrete image data, the spatially variant resolution of the logarithmic domain has to be considered. The sampling rate decreases rapidly towards the image margins so that an antialiasing filter must be provided.

4.2. Hartley Transform

The discrete Hartley transform (DHT) may be used in cases where data is in the real domain and phase information is not required to substitute the more commonly used DFT designed for complex data. The benefits of calculating power spectra and power cepstra using the DHT are about 50% less required memory and about 40% faster program execution, at no loss in accuracy [28]. Algorithms for fast Hartley transforms (FHT) are described elsewhere [29, 30].

The DHT was introduced as a sum of double-sided sine and cosine transforms [31, 32], corresponding to the continuous integral transform presented by Hartley [33] already in 1942. The separated DHT of an $M \times N$ image $f(x, y)$ is given to

$$H_s(u, v) = \sum_{x=0}^{M-1}\sum_{y=0}^{N-1} f(x, y)\mathrm{cas}(2\pi ux)\mathrm{cas}(2\pi vy) \tag{23}$$

and

$$f(x, y) = \frac{1}{MN}\sum_{u=0}^{M-1}\sum_{v=0}^{N-1} H_s(u, v) \times \mathrm{cas}(2\pi ux)\mathrm{cas}(2\pi vy) \tag{24}$$

denotes the inverse Hartley transform, with $\mathrm{cas}(\theta) = \cos(\theta) + \sin(\theta)$ being an abbreviation adopted from Hartley [33]. According to (23), the Fourier power spectrum is calculated exploiting the relationship between the DHT and the power spectrum [34],

$$|F(u, v)|^2 = \frac{1}{4}[H_s(-u, v) + H_s(u, -v)]^2 + \frac{1}{4}[H_s(u, v) - H_s(-u, -v)]^2 \tag{25}$$

In our registration approach, the Fourier power spectra used to calculate the power cepstrum (13), as well as the Fourier power spectra used for rotation detection (11), are numerically obtained with equation (25).

Fig. 4. *In vitro* object. This part of a human mandible was used to create the sequences shown in (Fig. 6). In addition, a Branemark screw was implanted in the interdental region to obtain the third sequence (Fig. 6 bottom). For image acquisition, the dry jawbone was mounted on the adjustment device (Fig. 5).

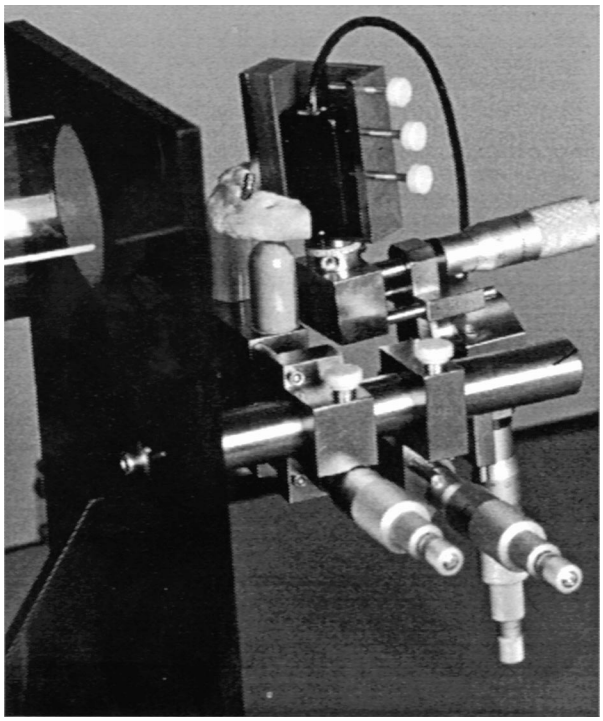

Fig. 5. Device for image acquisition. This figure shows the mechanical device constructed for the *in vitro* study. On the upper left, a part of the mounted X-ray tube is visible. The X-rays pass the phantom and were converted into a digital image with the aid of the CCD sensor, shown in the upper right. The sequences (Fig. 6) have been produced using the Sens-A-Ray system [11] (REGAM Medical Systems, Sundsvall, Sweden). The micrometer screws allow the adjustment of translations with a precision of 1/100 millimeter. The scales on the revolving axes can be adjusted in steps of one degree. Note that the *in vitro* object used in this study (Fig. 4) is different from the phantom shown here.

5. *In vitro* STUDY

At first, our method for generating test images with *a priori* known misalignment is described. Based on 855 pairs of dental radiographs, the three proposals for the rotation detection step are systematically compared next.

5.1. Test Image Acquisition

Taking into account the influence of real structures in the nonoverlapping regions of both medical images to be registered, we compared the three approaches for the rotation detection step by means of *in vitro* images and not on simulated data. High-resolution X-ray CCD-sensors are recently available in dental radiology [11, 12]. Therefore, our study bases on *in vitro* acquired digital radiographs of a human jaw (Fig. 4). A mechanical device was constructed fixing the positions of X-ray tube, phantom, and sensor (Fig. 5). The micrometer screws allow precise movements of sensor and object. Rotations can be adjusted on the revolving axes with an accuracy of one degree.

Three image sequences have been generated by equidistantly moving or rotating the sensor perpendicular to the central ray.

• The translation sequence containing 41 frames has been obtained moving the jawbone 40 mm perpendicular to the central ray along the 17.2 mm width sensor (Fig. 6 top).

• The combined sequence including 33 images has been acquired rotating the object with a total amount of 180° while the rotation axis was not in the center of the sensor. Therefore, this sequence contains not only rotational but also translational shifts (Fig. 6, middle).

• The third sequence has been produced similar to the second one. Additionally, the X-ray's dose was modified during the acquisition of the sequence. Therefore, the images differ in intensity and noise. The captured section of the dry mandible includes front teeth, premolars, and a dental implant produced by Branemark (Fig. 6, bottom).

Combining the sequence's single frames in pairs of two, altogether 855 different sets of X-rays with *a priori* known relative movement were registered applying all three techniques.

5.2. Results

Our extensive simulations proved the rotation detection step by means of the cross-correlation as well as the entropy measure to be superior to the one based on the cepstrum filter (table). For this, the entropy is calculated on the one-dimensional histogram of the difference image. The superiority of entropy and correlation to the cepstrum is because Fourier power spectra instead of real images are to be adjusted in the rotation detection step. Since correlation and entropy yield similar results, the entropy method is preferable due to its

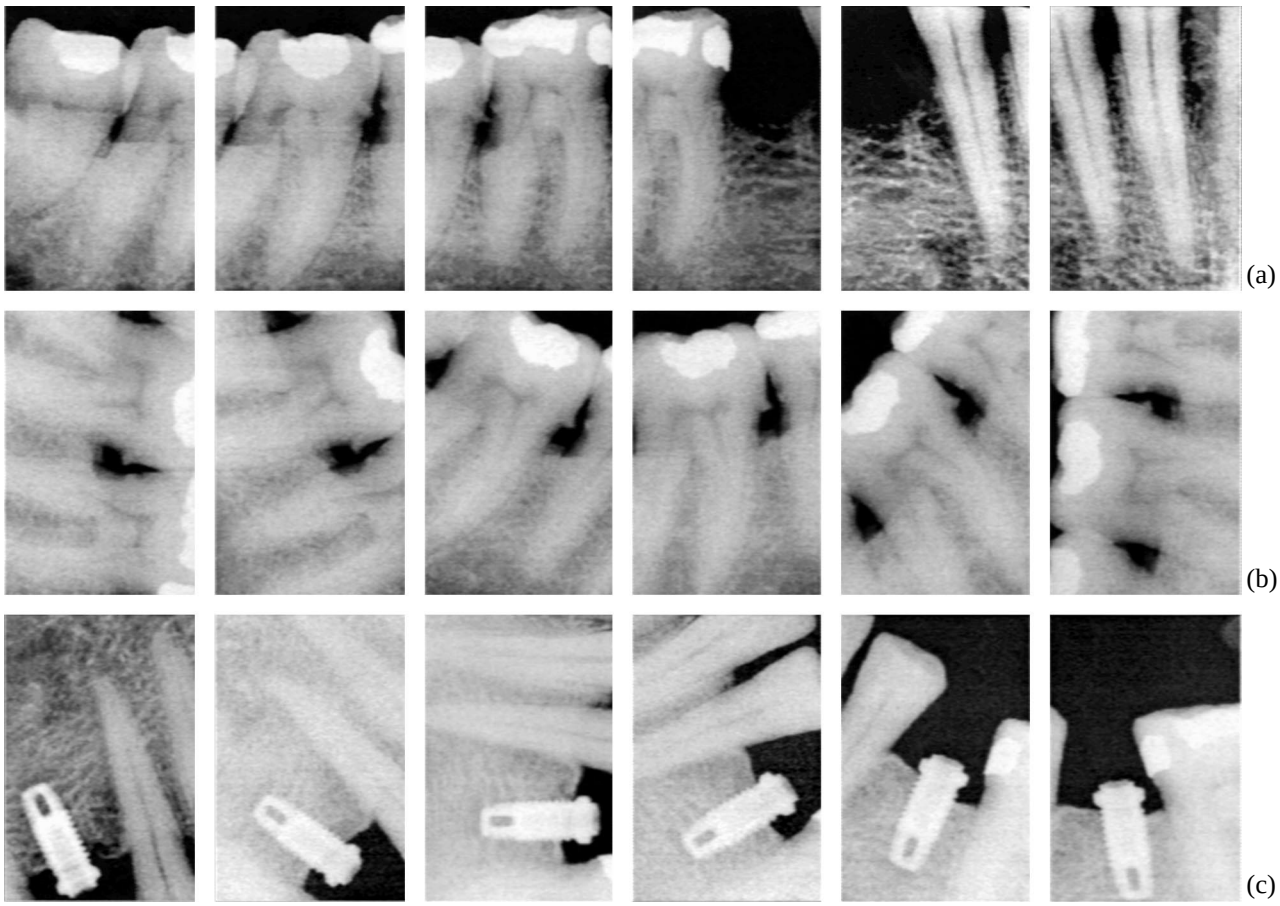

Fig. 6. Test image sequences. The lines at top, middle, and bottom show single frames equidistantly captured from the *in vitro* acquired test sequences: (a) pure translation; (b) translation and rotation; (c) translation, rotation, and intensity variation, respectively.

numerical efficiency. While the correlation has to be computed in a two-dimensional image plane, the entropy is calculated on a one-dimensional histogram function.

The union of the logarithmic polar mapped power spectra with the entropy-measure leads to a precise determination of rotational differences. One image is adequately rerotated applying 4×4 bicubic spline functions [22, 23]. Subsequently, translations are corrected using common cepstral techniques. This method allows reliable (probability >80%) registration of pictures overlapping more than 70% with rotational components less than 15 degrees, regardless of noise or intensity variations occurring in clinical radiographs. This result is deducted from the lines marked in table.

6. APPLICATION TO MEDICAL IMAGES

As we have pointed out before, the object/background model is invalid in most medical applications. The entire image must be regarded as the object. In addition, the object may be modified between the moments of capturing the two images to be spatially registered. These modifications can be classified in three categories: local differences, local and global deformations, and, of course, combinations of these classes.

The next paragraphs describe the application of the entropy-based rotation-extended cepstrum technique for each class of object modification.

6.1. Local Differences

If the spatial geometry of object and imaging system is constant or reproducible between the moments of image acquisition and hard tissue is imaged, only confined differences in the image may occur. In most cases these changes are of diagnostical interest and have to be detected or measured.

Examples are tuberculosis inspection in thorax X-rays or the determination of bone changes in dental implantology. In the latter case, implant patients are asked half-yearly for a recall examination with an intraoral X-ray taken routinely. The dentist has to compare radiographs differing in translation, rotation, intensity,

The table shows the results evaluating 855 pairs of dental radiographs with *a priori* known displacement. The three horizontal blocks refer to the three captured X-ray sequences (Fig. 6). The second column denotes the relative displacement of the two radiograps being registered. For the rotated sequences, the angle α of rotational movement is given. The third column denotes the overall number of image pairs evaluated with this movement. The block on the right shows the numbers and percentages of correct registrations using the three different methods to detect the rotational movement. The columns: Cepstrum, Correlation, and Entropy refer to equations (17), (19), and (21), respectively. The extensive simulations prove the rotation detection by means of cross-correlation or entropy to be superior to the one based on the cepstrum. The lines marked with an arrow (⟶) are referred to in the text judging the usefulness of the techniques.

Sequence	Displacement	Image pairs no.	Rotation detection performing					
			ceptrum		correlation		entropy	
			no.	%	no.	%	no.	%
Translation	0.00%	41	41	100	41	100	41	100
	5.81%	40	40	100	40	100	40	100
	11.63%	39	39	100	39	100	39	100
	17.44%	38	36	95	38	100	38	100
	23.26%	37	34	92	34	92	34	92
	⟶ 29.07%	36	26	72	29	80	30	83
	34.88%	35	20	57	21	60	21	60
	40.70%	34	11	32	15	44	17	50
	46.51%	33	6	18	5	15	7	21
Fixed Translation and Rotation	0.00°	33	33	100	33	100	33	100
	5.45°	32	29	88	32	100	32	100
	10.90°	31	18	58	31	100	31	100
	⟶16.35°	30	8	27	28	93	29	97
	21.80°	29	6	21	21	72	23	79
	27.25°	28	5	19	19	68	19	68
	32.70°	27	4	15	11	41	12	44
	38.15°	26	5	19	9	34	11	42
	43.60°	25	3	12	6	24	7	28
Fixed Translation and Rotation and Randomized Intensity Variation	0.00°	33	33	100	33	100	33	100
	5.45°	32	18	56	28	88	30	94
	10.90°	31	13	42	27	87	28	90
	⟶ 16.35°	30	9	30	22	73	24	80
	21.80°	29	7	24	16	55	19	66
	27.25°	28	2	7	9	32	13	46
	32.70°	27	4	15	11	40	13	48
	38.15°	26	1	4	4	18	9	34
	43.60°	25	0	0	4	16	6	24
Sum		855	451	52.75	606	70.88	639	74.74

and noise to assess the periimplant bone status, e.g., the two radiographs in the left part of (Fig. 1). The current image (Fig. 1, bottom left) is registrated to the reference image (Fig. 1, top left) using the new technique resulting in a rotation and a translation (Fig. 7, left). After cepstral filtering and contrast adjustment (Fig. 7, middle), the subtraction of both images (Fig. 7) proves the successful registration. Periimplant bone lesions are easily detectable in the subtraction image as dark spots. Note that there is also bone destruction in the interdental region. Therefore, the rotation-extended cepstrum technique has become the fundamental part of the automatical densitometric analysis of the alveolar bone structures in dental implantology.

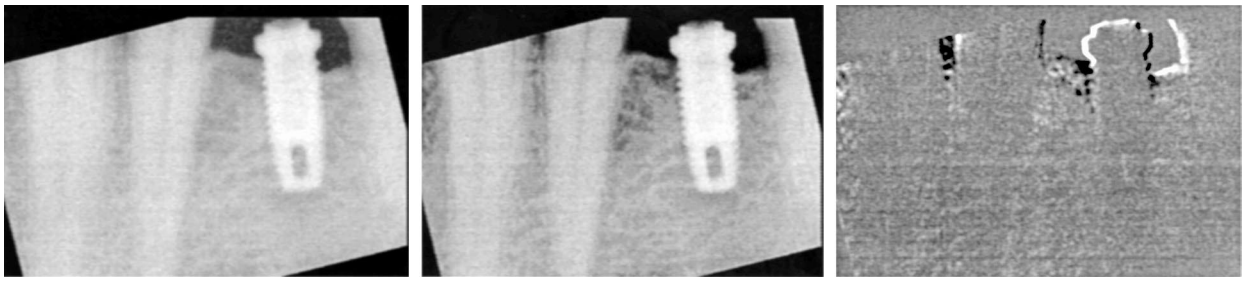

Fig. 7. Registration with local differences. The *in vitro* radiograph shown in (Fig. 1) on the bottom left was adjusted to the reference image (Fig. 1 top left) using the rotation-extended cepstrum technique. After geometrical (left) and contrast (middle) adjustment, the reference image was subtracted from the subsequent image (right). The simulated changes in bone structure are clearly visible as dark spots in the difference image. The black and white borders of the implant are artefacts caused by misalignment.

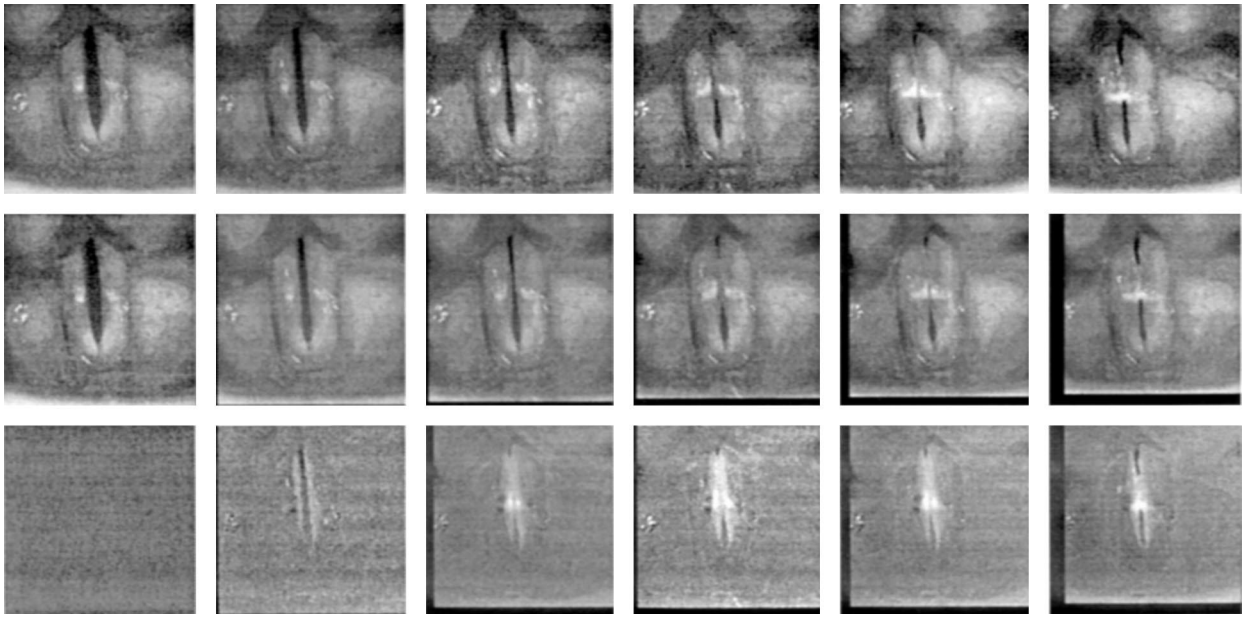

Fig. 8. Registration with local deformations. The top line shows single frames captured from a stroboscopic video sequence of moving vocal cord. The results of geometrical and contrast adjustment is shown in the middle and the difference images after registration are shown in the bottom line. The top left frame is used as reference for subtraction.

6.2. Local Deformations

The second class of object modifications are deformations occurring in some image regions. They are usually caused by muscle contraction *in vivo* imaging soft tissues. The endoscopic inspection of the larynx examples this category. Detuned triggering the stroboscopic illumination with the fundamental wave of the patient's phonation allows the assessment of the vocal cord's vibration. Considering single frames of the sequence, the vocal cords deform while the surrounding larynx is figured immovable.

The top line in (Fig. 8) shows six frames captured from a videolaryngoscopic image sequence. Additional to the vibration of the vocal cords, the patient and/or the examiner moves. The second line in (Fig. 8) displays the same frames after correction for patient movements using our new technique. The subtractions after contrast adjustment are arranged in the bottom line in (Fig. 8) with the first frame always used as reference. In the adjusted sequence, for example the area of vocal cord vibration is measurably opening new diagnostic benefits for the characterizing of laryngitis and chorditis.

6.3. Global Deformations

The last class of object movements in medical images are global deformations. The images to be adjusted are similar, but there are no image regions exactly corresponding. Therefore, this case is the most difficult one. It is not only complicated to register the images, but also to judge the quality of registration.

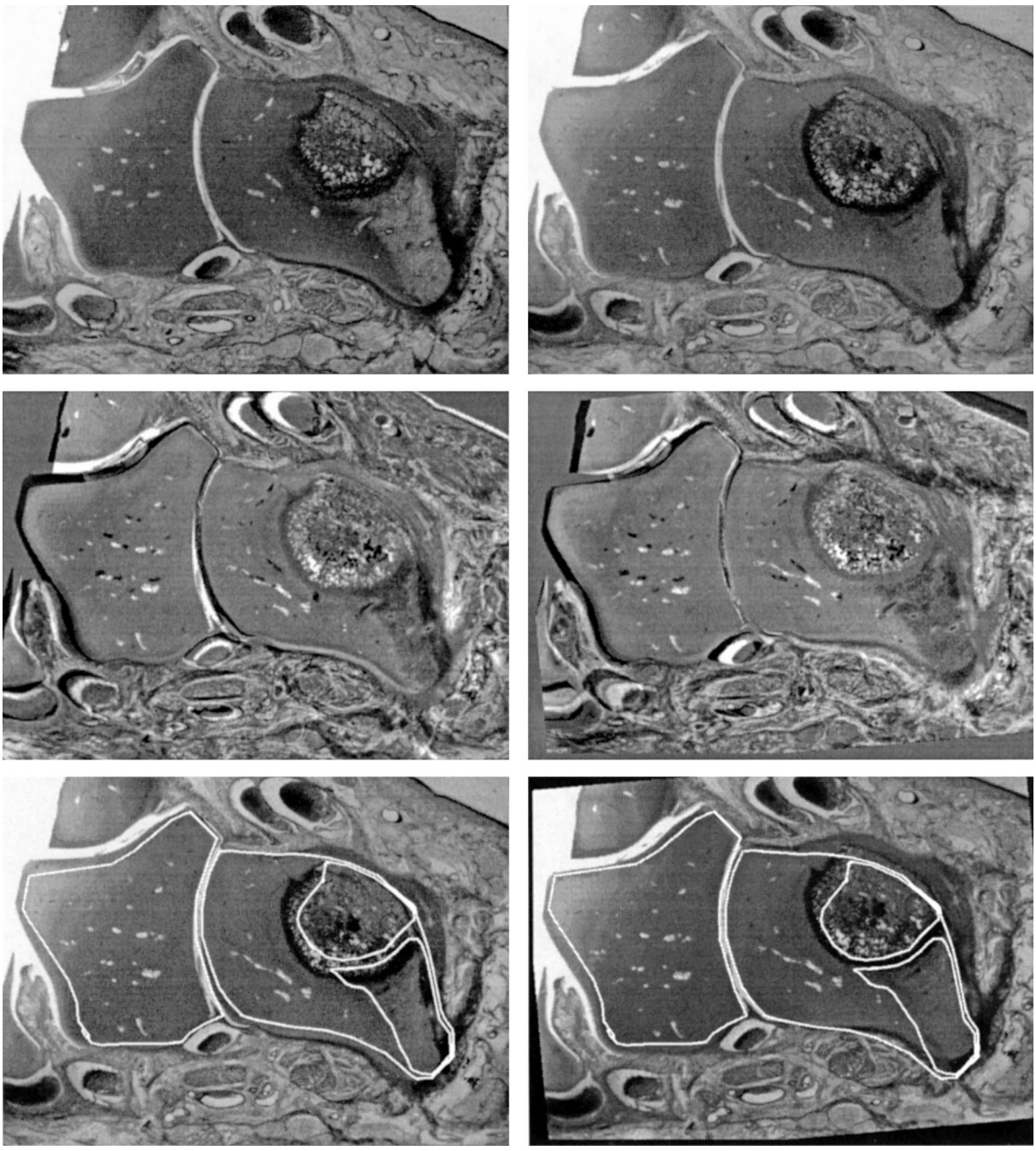

Fig. 9. Registration with global deformations. Two successive histological slices of a human ankle are shown in top line of figure. The difference images before and after adjustment are displayed in the middle line on the left and the right, respectively. The overlay of the outlines taken from image 1 to image 2 is shown in the bottom before and after registration of image 2, respectively.

Figure 9 shows two successive histological slices imaged from a human ankle. Translational and rotational distortions occurred during the subsequent digitization. Before the frames can be used to generate a three-dimensional reconstruction of the convex articular surface they have to be registered. The subtraction before and after registration using the new technique is less suitable to appraise the quality of alignment (Fig. 9, middle line). Although one might prefer the subtraction after registration, both difference images exhibit a lot of dominant structures in all image areas.

To evaluate the new technique registering images with global deformations, the outlines of bone and cartilage have been marked in the reference image by an expert. These lines are overlayed to the current image before and after registration (Fig. 9, bottom) proving the correct adjustment of both slices based on the rotation-extended cepstrum technique.

7. CONCLUSION

The common cepstrum technique was restricted so far to the detection of pure translational shifts. Combining the entropy analysis with the Fourier–Mellin RST-invariant image descriptor, cepstrum is now extended to rotations as well. Is was proven that the new algorithm allows reliable registration of images displaced less than 30% with rotations smaller than 15 degrees, regardless of noise or intensity variations.

The proposed method was successfully applied in several cases of medical imaging. The registration of images differing locally was evaluated on radiographs taken in dental implantology. Although local occurring deformations are present, the frames of video sequences of the vocal cords are precise adjusted. The capacity of our method to register images with global deformations is proven based on histological slices of human ankles. All categories of differences occurring in the two images to be registered, as well as noise and intensity variations, cannot prohibit successful registration.

Therefore, the new technique is applied as the initial image processing step for the densitometric analysis of bone lesions in dental radiographs. This initial registration permits the subtraction of sequential images followed by an automatic segmentation of bone lesions and measurement of lesion areas. In dental radiology, radiographs are taken from the same dental region of the same patient at different points of time. Translational as well as rotational misalignments can be precisely compensated *in vivo* with the new method even if other distortions of the projective geometry occur.

8. ACKNOWLEDGMENTS

This research was undertaken in the project Freehand Subtraction Radiography. The support of the German Research Community grant DFG Re 427/5-1 is gratefully acknowledged. The authors would like to thank Ulrich Breuer, Medical University of Lübeck, Germany, for providing us with the histological slices and Ingrid Scholl, The Aachen University of Technology, Germany, for intensive discussions and comments.

REFERENCES

1. Lehmann, T., Goerke, C., Kaupp, A., and Repges, R., Cepstrum Analysis Extended to Rotational and Translational Registration, *Proc. of the Third Int. Conf. on Pattern Recognition and Information Analysis (PRIA'95)*, Ablameyko, S. and Soldek, J., Eds., Szczecin: Zapol, 1995, vol. 3, pp. 111–115.
2. Cideciyan, A.V., Registration of Ocular Fundus Images, *IEEE Magazine EMB*, 1995, vol. EMB14(1), pp. 52–58.
3. Chen, C.S., Defrise, M., and Deconinck, F., Symmetric Phase-Only Matched Filtering of Fourier–Mellin Transforms for Image Registration and Recognition, *IEEE Trans. on PAMI*, 1994, vol. PAMI-16(12), pp. 1156–1168.
4. Pratt, W.K., *Digital Image Processing*, New York: Wiley, 1978.
5. Reiss, T.H., *Recognizing Planar Objects Using Invariant Image Features*, Berlin: Springer, 1993.
6. van der Stelt, P.F., Ruttimann, U.E., and Webber, R.L., Determination of Projections for Subtraction Radiography Based on Image Similarity Measurements, *Dentomaxillofacial Radiology*, 1989, vol. 18, pp. 113–117.
7. Dunn, S.M., van der Stelt, P.F., Fenesy, K., and Shah, S., A Comparison of Two Registration Techniques for Digital Cubtraction Radiography, *Dentomaxillofacial Radiology*, 1993, vol. 22, pp. 77–80.
8. Venot, A., Lebruchec, J.F., and Roucayrol, J.C., A New Class of Similarity Measures for Robust Image Registration, *Comput. Vision, Graphics, Image Process.*, 1984, vol. 28, pp. 176–184.
9. Venot, A. and Leclerc, V., Automated Correlation of Patient Motion and Grey Value Prior to Subtraction in Digitized Angiography, *IEEE Trans. on MI*, 1984, vol. MI-3, pp. 179–186.
10. Lee, D.J., Mitra, S., and Krile, T.F., Analysis of Sequential Complex Images, Using Feature Extraction and Two-Dimensional Cepstrum Techniques, *J. Opt. Soc. Am. A*, 1989, vol. 6, pp. 863–870.
11. Welander, U., Nelvig, P., Tronje, G., McDavid, W.D., Dove, S.B., Morner, A.C., and Cederlund, T., Basic Technical Properties of a System for Direct Acquisition of Digital Intraoral Radiographs, *Oral Surgery Oral Medicine Oral Pathology*, 1993, vol. 75, no. 4, pp. 506–516.
12. Hillen, W., Bockemühl, J., and Schmitt, W., Signal-to-Noise Performance of a Digital Dental X-ray Detector, *Proc. CAR'95, Computer Assisted Radiology*, Lemke, H.U., Inamura, K., Jaffe, C.C., and Vannier, M.W., Eds., Berlin: Springer, 1995, pp. 990–995.
13. Bracewell, R.N., *The Fourier Transform and Its Applications*, New York: McGraw-Hill, 1986, 2nd ed.
14. Colombo, S. and Lavoine, J., *Transformations de Laplace et de Mellin: Formulaires, Mode D'Utilisation*, Paris: Gauthier Villars, 1972.
15. Oberhettinger, F., *Tables of Mellin Transforms*, Berlin: Springer, 1974.
16. Schalkoff, R.J., *Digital Image Processing and Computer Vision*, New York: Wiley, 1989.
17. Haken, H., *Synergetic Computers and Cognition: A Top Down Approach to Neural Nets*, Berlin: Springer, 1991.
18. Lehmann, T., Goerke, C., Schmitt, W., and Repges, R., Rotations- und Translationsbestimmung durch eine erweiterte Kepstrum Technik, in *Mustererkennung 1995. Verstehen akustischer und visueller Information: Reihe Informatik aktuell*, Sagerer, G., Posch, S., and Kummert, F., Eds., Berlin: Springer, 1995, pp. 395–402.
19. Bogert, B.P., Healy, M.J.R., and Tukey, J.W., The Frequency Alanysis of Time Series for Echoes: Cepstrum, Pseudo-Autovariance, Cross-Cepstrum and Saphe Cracking, *Proc. of the Symp. of Time Series Analysis*, 1963, pp. 209–242.
20. Childers, D.G., Skinner, D.P., and Kemerait, R.C., The Cepstrum: A Guide to Processing, *Proc. of the IEEE*, 1977, vol. 65, pp. 1428–1443.
21. Oppenheim, A.V. and Schafer, R.W., *Digital Signal Processing*, Englewood Cliffs: Prentice-Hall, 1975.

22. Danielsson, P.E. and Hammerin, M., *High Accuracy Rotation of Images, Report No. LiTH-ISY I-1152*, Linkoping: Department of Electrical Engineering, Linkoping Univ., Sweden, 1990.
23. Danielsson, P.E. and Hammerin, M., *High Accurate Rotation of Images*, *CVGIP: Graphical Models and Image Processing*, 1992, vol. 54, pp. 340–344.
24. Wenzel, A., Effect of Manual Compared with Reference Point Superimposition on Image Quality in Digital Subtraction Radiography, *Dentomaxillofacial Radiology*, 1989, vol. 18, pp. 145–150.
25. Kapur, J.N., Sahoo, P.K., Wong, A.K.C., A New Method for Gray-Level Picture Thresholding Using the Entropy of the Histogram, *Computer Vision, Graphics, and Image Processing*, 1985, vol. 29, pp. 273–285.
26. Lehmann, T., Schmitt, W., Repges, R., and Sovakar, A., Mathematical Quality Standards for the Digital Freehand Subtraction Radiography, *Dentomaxillofacial Radiology*, 1995, vol. 24, p. 98.
27. Harris, F.J., On the Use of Windows for Harmonic Analysis with the Discrete Fourier Transform, *Proc. of the IEEE*, 1978, vol. 66, pp. 51–83.
28. Steckner, M.C. and Drost, D.J., Fast Cepstrum Analysis Using the Hartley Transform, *IEEE Trans. on ASSP*, 1989, vol. ASSP-37, pp. 1300–1302.
29. Bracewell, R.N., The Fast Hartley Transform, *Proc. of the IEEE*, 1984, vol. 72, pp. 1010–1018.
30. Besslich, P.W. and Lu, T., *Diskrete Orthogonaltransformationen*, Berlin: Springer, 1990.
31. Bracewell, R.N., The Discrete Hartley Transform, *J. Opt. Soc. Am. A*, 1983, vol. 73, pp. 1832–1835.
32. Bracewell, R.N., *The Hartley Transform*, New York: Oxford University Press, 1986.
33. Hartley, R.V.L., A more Symmetrical Fourier Analysis Applied to Transmission Problems, *Proc. IRE*, 1942, vol. 30, pp. 144–150.
34. Watson, A.B. and Poirson A, Separable Two-Dimensional Discrete Hartley-Transform, *J. Opt. Soc. Am. A*, 1986, vol. 3, pp. 2001–2004.

Thomas Martin Lehmann. Born in 1966. Received the Master degree in electrical engineering from the Aachen University of Technology in 1992. In 1992 he worked at the Institute for Measurement Technology, Aachen University of Technology. Since 1992 he has been with the Institute of Medical Informatics and Biometry, Aachen University of Technology. He is now the head of the medical image processing group. His research interests are in medical image processing and its applications to computer aided diagnosis. He has five publications and he is member of the International Association of Dento-Maxillo-Facial Radiology (IADMFR), the Institute of Electrical and Electronic Engineering (IEEE), and the International Society of Optical Engineering (SPIE). In 1993 he was given the DAGM-Preis'93. The award from the German Association for Pattern Recognition (DAGM) was given for his work on automatical strabometry using Hough transformation and covariance filtering.