ULTRA-FAST HOLOGRAPHIC RECORDING AND AUTOMATIC 3D SCAN MATCHING OF LIVING HUMAN FACES

DOMINIK GIEL, SUSANNE FREY, ANDREA THELEN,
JENS BONGARTZ, PETER HERING
caesar foundation
Ludwig Erhard Allee 2, D-53175 Bonn, Germany
E-mail: {giel,frey,thelen,bongartz,hering}@caesar.de

ANDREAS NÜCHTER, HARTMUT SURMANN,
KAI LINGEMANN, JOACHIM HERTZBERG
Fraunhofer Institute for Autonomous Intelligent Systems (AIS)
Schloss Birlinghoven, D-53754 Sankt Augustin, Germany
E-mail: {nuechter,surmann,lingemann,hertzberg}@ais.fraunhofer.de

3D models of the skin surface of patients are created by ultra-fast holography and automatic scan matching of synchronously recorded holograms. By recording with a pulsed laser and continuous-wave optical reconstruction of the holographic real image, motion artifacts are eliminated. Focal analysis of the real image yields a surface relief of the patient. To generate a complete 360° patient model, several synchronously recorded reliefs are registered by automatic scan matching. We find the transformation consisting of a rotation and a translation that minimizes a cost function containing the Euclidian distances between points pairs from two surface relief maps. A variant of the ICP (Iterative Closest Points) algorithm[2] is used to compute such a minimum. We propose a new fast approximation based on kD-trees for the problem of creating the closest point pairs on which the ICP algorithm spends most of its time.

1. Introduction

To treat diseases, injuries and congenital or acquired deformities of the head and neck, maxillo-facial surgeons deal with complex surgery. For example, the correction of disfiguring facial birth defects requires the manipulation of scull bones with maximum precision. The pre-operative simulation of such procedures requires a 3D computer model of the patient's face. We describe an approach to create such a 3D patient model by ultra-fast holographic recording and automatic scan matching of synchronously captured holograms. The pulsed hologram records the patient's portrait within a single laser shot (pulse duration appr. 35 ns). This so-called master-hologram contains the complete 3D spatial information which, due to the extremely short recording time, is not affected by involuntary patient movements.

In a second step, the real image of the hologram is optically reconstructed with a cw-laser. By moving a diffusor-screen through the real image, a series of 2D images is projected and digitized with a CCD camera. This process is referred to as hologram tomography[3]. Each projection shows the surface contour of the patient where the image is in focus. The method was first introduced as the *locus of focus* technique[8] in the context of non-medical imaging. Beside the desired intensity from in-focus points from the object contour, each captured image also contains a blurred background of defocused parts of the real image. The main problem of locating the surface is therefore to distinguish between focused and unfocused regions in each slice. This procedure yields a relief map of the visible (as seen from the hologram) parts of the patient. In order to record a complete 360° model of a patient, multiple holograms are recorded synchronously, i.e. with the same laser pulse.

Subsequently, the resulting relief maps are registered (i.e. their relative orientation is calculated) by automated scan matching. The problem of automated scan matching is to find a transformation, consisting of a rotation and a translation, that minimizes a cost function that contains the Euclidian distances between points pairs[6] which both represent the same surface shape. Given that the surface shapes are acquired independently from locus-of-focus analysis of two synchronously recorded holograms, such an approach yields 360° models of complex surfaces.

2. Hologram tomography

2.1. *Recording and optical reconstruction*

Portrait holograms are recorded with a holographic camera of the type GP-2J (brand Geola) with master-oscillator and second harmonic generation. The resulting wavelength of 526.5 nm has a small penetration depth into skin to minimize light diffusion. Careful mode selection leads to a coherence length of approximately 6 m. The laser output is split into three beams: Two of them serve for homogeneous illumination of the object. They are expanded by concave lenses and diffusor plates at the output ports of the laser. The third beam serves as reference beam. The hologram plate (30 cm × 40 cm, VRP-M emulsion by Slavich) is developed with SM-6 and bleached with PBU-Amidol to obtain phase holograms. A frequency doubled cw Nd:YAG laser (COHERENT Verdi V-2) is used to reconstruct the holographic real image. To obtain the 2D projections of the real image, a diffusor (diffusor thickness 40 m, diameter 380 mm) is moved on a computer controlled linear positioning stage (PI M-531.DD, max. resolution 10 m) through the image volume. The diffusor images are digitized by a KODAK Megaplus ES 4.0 digital camera with 2048 x 2048 pixels.

2.2. *Locus of focus*

To analyze the sequence of 2D-projections, the so-called slices, we use digital image processing. As an approximation we assume that the surface derived from a single hologram has no undercuts. Therefore there can be no two surface points with the same (x, y)-coordinate and the surface can be represented by a relief map. As already mentioned, each captured slice contains the specific focused information representing the object shape contour and a defocused background. The task is thus to distinguish between focused and defocused image regions: To evaluate the sharpness of an image in an conventional imaging system, several algorithms have been proposed in the field of image processing. We found that the best measure for image sharpness is the statistical variance $V_{(x,y)}$ of the light intensity on pixel adjacent to (x, y). For each lateral coordinate (x, y), the sharpness measure $V_{(x,y)}(z)$ is a positive, real number. The axial coordinate $z_{(x,y)}$ is assigned by choosing $z_{(x,y)}$ to satisfy $V_{(x,y)}(z_{(x,y)}) \geq V_{(x,y)}(z) \quad \forall \quad z$. Thus each holographic real image gives a relief map of the object surface.

3. Automatic 3D Scan Matching

Since the relative orientation of multiple reliefs is usually not known a priori in the desired accuracy, these reliefs have to be merged in one coordinate system. This process -also known as *scan matching* since it originally referred to the orientation of *scans* from laser triangulation systems- is called registration. The geometric structure of overlapping 3D reliefs that correspond to a single shape has to be considered for registration. In general, scan matching approaches can be classified into two categories:

(1) Matching as an optimization problem uses a cost function to evaluate the quality of the 3D scan alignment. The range images are registered by determining the rigid transformation (rotation and translation) which minimizes the cost function.
(2) Feature based matching extracts distinguishing features of the range images and uses corresponding features for calculating the alignment the reliefs.

3.1. *Matching as an Optimization Problem*

The following method of registration of point sets is part of many publications so only a short summery is given here. The complete algorithm was published 1992 first and can be found, e.g., in[2]. The method is called Iterative Closest Points (ICP) algorithm.

Given two independently derived set of 3D points, M (model set, $|M| = N_m$) and D (data set, $|D| = N_d$), which correspond to a single shape, we want to find the transformation consisting of a rotation $\mathbf{R}$ and a translation $\mathbf{t}$ which minimizes the following cost function:

$$E(\mathbf{R}, \mathbf{t}) = \sum_{i=1}^{N_m} \sum_{j=1}^{N_d} w_{i,j} \left|\left|\mathbf{m}_i - (\mathbf{R}\mathbf{d}_j + \mathbf{t})\right|\right|^2 . \quad (1)$$

The value of 1 is assigned to $w_{i,j}$ if the i-th point of set M describes the same point in space as the j-th point of set D. Otherwise $w_{i,j}$ is set to 0. Two things have to be calculated: First the corresponding points and second the transformation $\mathbf{R}$ and $\mathbf{t}$ that minimizes $E(\mathbf{R}, \mathbf{t})$ using the point correspondeces.

The ICP calculates iteratively the point correspondences. In each iteration step the algorithm selects the closest points as correspondences and calculates the transformation $(\mathbf{R}, \mathbf{t})$ for Eq. (1). It is shown that the iteration terminates in a (local) minima[2]. The assumption is that in the last iteration step the point correspondences are correct.

In each iteration the transformation is calculated by the quaternion based method of Horn[5]. A unit quaternion is a 4 vector $\dot{q} = (q_0, q_x, q_y,_z)^T$, where $q_0 \geq 0, q_0^2 + q_x^2 + q_y^2 + q_z^2 = 1$. It describes a rotation axis and an angle to rotate around that axis. A 3×3 rotation matrix $\mathbf{R}$ is calculated from the unit quaternion according the the following scheme:

$$\mathbf{R} = \begin{pmatrix} (q_0^2 + q_x^2 - q_y^2 - q_z^2) & 2(q_x q_y + q_z q_0) & 2(q_x q_z + q_y q_0) \\ 2(q_x q_y + q_z q_0) & (q_0^2 - q_x^2 + q_y^2 - q_z^2) & 2(q_y q_z - q_x q_0) \\ 2(q_z q_x - q_y q_0) & 2(q_z q_y + q_x q_0) & (q_0^2 - q_x^2 - q_y^2 + q_z^2) \end{pmatrix} .$$

To determine the transformation, the mean values (centroid vectors) $\mathbf{c}_m$ and $\mathbf{c}_d$ of the points that contribute to the matching are subtracted from all points in M and D respectively, resulting in sets M' and D'. The rotation expressed as quaternion that minimizes equation (1) is the largest eigenvalue of the cross-covariance matrix

$$\mathbf{N} = \begin{pmatrix} (S_{xx} + S_{yy} + S_{zz}) & (S_{yz} + S_{zy}) & (S_{zx} + S_{xz}) & (S_{xy} + S_{yx}) \\ (S_{yz} + S_{zy}) & (S_{xx} - S_{yy} - S_{zz}) & (S_{xy} + S_{yx}) & (S_{zx} + S_{xz}) \\ (S_{zx} + S_{xz}) & (S_{xy} + S_{yx}) & (-S_{xx} + S_{yy} - S_{zz}) & (S_{yz} + S_{zy}) \\ (S_{xy} + S_{yx}) & (S_{yz} + S_{zy}) & (S_{zx} + S_{xz}) & (-S_{xx} - S_{yy} + S_{zz}) \end{pmatrix} ,$$

with $S_{xx} = \sum_{i=1}^{N_m} \sum_{j=1}^{N_d} w_{i,j}\ m'_{ix} d'_{jx}$, $S_{xy} = \sum_{i=1}^{N_m} \sum_{j=1}^{N_d} w_{i,j}\ m'_{ix} d'_{jy}$, $\ldots$. After the calculation of the rotation $\mathbf{R}$ the translation is $\mathbf{t} = \mathbf{c}_m - \mathbf{R}\mathbf{c}_d$ [5]. Fig. 1 shows three steps of the ICP algorithm[a], The corresponding surface meshes are given in Fig. 2.

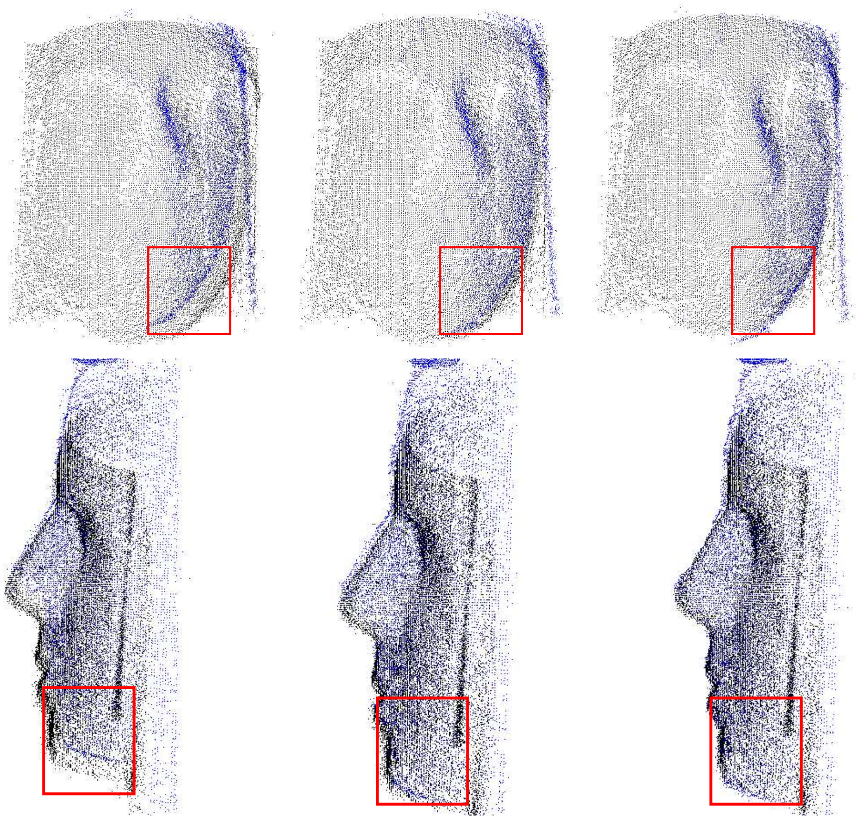

Figure 1. Registration of two 3D reliefs with the ICP algorithms. Left: Initial alignment. Middle: Alignment after 4 iterations. Right: Final alignment after 85 iterations.

3.2. *Time Complexity Reduction*

The ICP algorithms spends most of its time in creating the point pairs. kD-trees (here $k = 3$) have been suggested for speed up the data access[1].

[a]For an animation of this result please refer to the following website: `http://www.ais.fraunhofer.de/face`.

They are a binary tree with terminal buckets. The data is stored in the buckets, the keys are selected, such that a data space is divided into two equal parts. This ensures that a data point can be selected in $O(\log n)$ at average.

Recently, Greenspan and Yurick have introduced approximate kd-trees (Apr-kd-tree)[4]. The idea behind this is to return as an approximate nearest neighbor $\mathbf{p}_a$ the closest point $\mathbf{p}_b$ in the bucket region where the given point $\mathbf{p}$ lies. This value is determined from the depth-first search, thus expensive Ball-Within-Bounds tests and backtracking are not necessary[4]. In addition to these ideas we avoid the linear search within the bucket. During the computation of the Apr-kd-tree the mean values of the points within a bucket are computed and stored. Then the mean value of the bucket is used as the approximate nearest neighbor, replacing the linear search. Table 1 summarizes the results.

Table 1. Computing time and number of ICP iterations to align two 3D reliefs (Pentium-IV-2400).

point pairing method	time	# ICP iterations
brute force search	4 h 25 min	87
kD–tree	14.2 sec	87
Apx-kD–tree	10.9 sec	85

3.3. *Matching Multiple 3D Reliefs*

To digitalize human faces without occlusions, multiple depth maps have to be registered. After registration the scene has to be globally consistent. A straightforward method for aligning several 3D reliefs is *pairwise matching*, i.e., the new scan is registered against the scan with the largest overlapping areas. The latter one is determined in a preprocessing step. An alternative method is *incremental matching*, i.e., the new 3D relief is registered against a so called metascan, which is the union of the previous acquired and registered reliefs. Each scan matching has a limited precision. Both methods accumulate the registration errors such that the registration of many reliefs leads to inconsistent scenes[6].

Pulli presents a registration method that minimizes the global error and avoids inconsistent scenes[7]. Based on the idea of Pulli we designed a method called *simultaneous matching*. Hereby the first scan is the masterscan and determines the coordinate system. This scan is fixed. The following steps register all reliefs and minimize the global error:

(1) Based on the prior knowledge about the relative coordinate systems, that needs not be precise or complete, pairwise matching is used to find a start registration for a new scan. This step speeds up computation.
(2) A queue is initialized with the new scan.
(3) Three steps are repeated until the queue is empty:
    (a) The current scan is the first scan of the queue. This scan is removed from the queue.
    (b) If the current scan is not the master scan, a set of neighbors (set of all reliefs which overlap with the current scan) is calculated. This set of neighbors form the point set M. The current scan forms the data point set D and is aligned with the ICP algorithms.
    (c) If the current scan changes its location by applying the transformation, then each single scan of the set of neighbors, which is not in the queue is added to the end of the queue.

4. Results and Conclusions

We have demonstrated that the techniques of hologram tomography and automated scan matching can be combined to create 360° models of living human heads. Due to the ultra-fast acquisition, these models are inherently free of motion artifacts as opposed to surfaces models recorded with laser triangulation. The ICP algorithm is thus a valuable tool to register high-resolution models of the living human skin surface which are today commonly recorded by laser scanning and limited by the low acquisition speed in conjunction with the motion of breathing, heartbeat and involuntary movements of the patient. Even with relief models from a single hologram, holographic recordings are of value for the documentation and prediction of complex, maxillo-facial surgery. We expect that the possibility to assemble multiple holographic reconstructions into accurate 3D patient models will greatly increase the acceptance of the surface models obtained by hologram tomography.

Acknowledgment

The authors wish to acknowledge the cooperation with the medical groups of Prof. Dr. Dr. H.F. Zeilhofer (Kantonsspital Basel, Switzerland) and Prof. Dr. Dr. C.U. Fritzemeier (Universitätsklinikum Düsseldorf, Germany).

Figure 2. Surface meshes. Left: 3D relief as surface mesh. Right: Registed 2nd mesh (blue) merged with the first one.

References

1. J. L. Bentley. Multidimensional binary search trees used for associative searching. *Communications of the ACM*, 18(9):509 – 517, September 1975.
2. P. Besl and N. McKay. A method for Registration of 3–D Shapes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14(2):239 – 256, February 1992.
3. D. M. Giel. *Hologram tomography for surface topometry*. PhD thesis, Mathematisch-Naturwissenschaftliche Fakultät der Heinrich-Heine-Universität Düsseldorf, 2003.
4. M. Greenspan and M. Yurick. Approximate K-D Tree Search for Efficient ICP. In *Proceedings of the 4th IEEE International Conference on Recent Advances in 3D Digital Imaging and Modeling (3DIM '03)*, pages 442 – 448, Banff, Canada, October 2003.
5. B. Horn. Closed–form solution of absolute orientation using unit quaternions. *Journal of the Optical Society of America A*, 4(4):629 – 642, April 1987.
6. A. Nüchter, H. Surmann, K. Lingemann, and J. Hertzberg. Consistent 3D Model Construction with Autonomous Mobile Robots. In *Proceedings of the KI 2003: Advances in Artificial Intelligence. 26th Annual German Conference on AI, Proceedings Springer LNAI vol. 2821*, pages 550 – 564, Hamburg, Germany, September 2003.
7. K. Pulli. Multiview Registration for Large Data Sets. In *Proceedings of the 2nd International Conference on 3D Digital Imaging and Modeling (3DIM '99)*, pages 160 – 168, Ottawa, Canada, October 1999.
8. K. A. Stetson. Holographic surface contouring by limited depth of focus. *Applied Optics*, 7(5):987 – 989, 1968.