Registration of PE Segment contour deformations in digital High-Speed Videos
– submitted to Medical Image Analysis –

Michael Stiglmayr∗† Raphael Schwarz‡ Kathrin Klamroth†
Günter Leugering‡ Jörg Lohscheller‡

Abstract
Oncologic therapy of laryngeal cancer may necessitate a total excision of the larynx which results into loss of voice. Voice rehabilitation can be archived using mucosal tissue vibrations at the upper part of the esophagus which serve as substitute voice generating element (PE segment). The quality of the substitute voice is closely related to vibratory characteristics of the PE segment. By means of a high-speed camera the dynamics of the PE segment can be recorded in real-time. Using image processing the deformations of the PE segment are extracted from the image series as deforming contours. Commonly, the characterization of PE dynamics bases on the spectral analysis of the time varying contour area. However, this constitutes an integral approach which masks most of the specific dynamics of PE deformations.

We present an algorithm that automatically registers one segmented frame of the video sequence to the next frame to derive discrete 2-D trajectories of PE vibrations. By concatenation of the obtained transformations this approach provides a total registration of PE segment contours. We suggest a mixed-integer programming formulation for the problem that combines an advanced outlier handling with the introduction of dummy points in regions that newly open up, and that includes normal information in the objective function to avoid unwanted deformations. Numerical experiments show that the implemented alternate convex search algorithm produces robust results demonstrated in two high speed recordings of laryngectomee subjects.

∗Corresponding author. E-mail: stiglmayr@am.uni-erlangen.de; Tel: +49 9131 8528306; FAX: +49 9131 8528126; Address: Institute of Applied Mathematics, Martensstr. 3, 91058 Erlangen, Germany; http://www2.am.uni-erlangen.de/~stiglmayr
†Institute of Applied Mathematics, University of Erlangen – Nürnberg, partially supported by DFG, Sonderforschungsbereich 603 subproject C11
‡Department of Phoniatrics and Pediatric Audiology, University of Erlangen – Nürnberg, supported by DFG, Sonderforschungsbereich 603 subproject B5
1 Introduction

Depending on the stage and location of laryngeal cancer the surgical excision of the larynx may be the only adequate oncologic therapy. At a so-called total laryngectomy the malignancy is excised in conjunction with the larynx which requires a separation of the trachea and the esophagus to prevent uncontrolled mixing of breathing and swallowing Blom (2000). While the connection of the upper esophageal sphincter with the pharynx is not affected, the upper end of the trachea is sutured into the anterior skin of the neck. The artificial tracheal opening at the neck is called tracheostoma and facilitates the vitally important maintenance of respiration. Due to the disconnection of the trachea from the pharynx and the loss of the larynx a total laryngectomy results into the loss of voice which deeply affects the post-surgical social integration Eadie and Doyle (2004).

Rehabilitation of voice can be achieved by exciting tissue vibrations within the esophagus serving as substitute voice generator to compensate the function of the excised larynx. It is achieved by means of a silicon shunt valve that reconnects the trachea to the oral cavity. The prothesis builds up a unidirectional connection and permits air to pass from the trachea into the esophagus but prevents food from being aspirated into the trachea Blom (2000). Fig. 1.1 shows schematically the main principle of tracheoesophageal voice production. The shunt valve redirects the airstream into the esophagus when the tracheostoma is manually occluded during exhalation. At the upper part of the esophagus the streaming air excites tissue vibrations including the scar between esophagus and the pharynx which in turn modulate the airstream and generate the substitute voice signal. On account of the location of the oscillating tissue the sound emitting part of the esophagus is named pharyngo-esophageal segment, abbreviated PE segment. Finally as in normal voice production, the acoustical signal reaches the vocal tract, gets filtered by its resonance cavity, and is emitted as voice signal through the mouth Titze (1994).

In comparison with normal voice, the quality of tracheoesophageal voice is significantly reduced which affects the quality of life of patients which underwent a laryngectomy van As-Brooks et al. (2005). Thus, voice therapy plays a major part within the oncologic rehabilitation following surgery Schuster et al. (2004). Improvement of voice quality requires a comprehensive understanding of the substitute voice generating process. The quality of tracheoesophageal voice depends essentially on the vibration pattern of the PE segment. In order to investigate the relationship between PE vibrations and voice quality the vibrations of the PE segment have to be analyzed during sound production (i.e., phonation). For this purpose PE segment vibrations are investigated using endoscopic high-speed video systems which enable an insight into its anatomical, morphological, and dynamic characteristics in real-time. Current high-speed camera systems have a frame rate of 4000 Hz and a spatial resolution of 256 × 256 pixel (High-Speed Endocam, Wolf Corp., Knittlingen, Germany). Simultaneously, the emitted acoustic signal is recorded (microphone: type B&K 4129 Brüel & Kjaer Corp., sampling rate 44.1 kHz, discretization 16 bit).
A schematic draft of the examination situation is shown in Fig. 1.1. The patients are instructed to articulate a vowel /a/ in a comfortable way. The vibrations of the PE segment are recorded from a top-view position with a high-speed camera system which is coupled to a rigid endoscope. Fig. 1.2 shows short sequences of two high-speed recordings (A,B).

Within the sequences the deformations of the PE segment are captured. In each frame the mucosal tissue of the PE segment forms a constriction within the esophagus which constitutes the border of the enclosed opening. The visible opening formed by the PE segment is called pseudoglottis. The two high-speed movies show that the movements of the PE segment follow a quasi-periodic vibration pattern. However, the PE segment oscillations of different subjects diverge with respect to the shape, size, frequency, and orientation of pseudoglottal deformations.

For the extraction of PE dynamics an image processing procedure had been developed which bases on active contour models using an adapted gradient vector flow field (GVF-field) as external force Lohscheller et al. (2004, 2003). Fig. 1.3 shows the segmentation results of the high-speed movies A and B. The contours track the deformation of the pseudoglottis and confirm the quasi-periodic vibration characteristic of the PE segment.

By setting the spectral information of the pseudoglottal deformations into relation to the spectral properties of the emitted acoustical signal the interrelation between the substitute voice and the PE vibration could be verified Lohscheller.
et al. (2004); Schuster et al. (2005). Up to now, just the time-varying pseudoglottal area function is derived from the segmentation results and serves as measure to describe PE segment oscillations. As the analysis of the pseudoglottal area is an integral approach, the specific information of the properties of the PE tissue vibrations as oscillation directions, velocities and amplitudes gets lost. Also no predictions can be made how to alter the PE morphology by surgical intervention to modify the PE vibration pattern to improve voice quality.

In this work we propose a novel approach to register consecutive contours of the PE segment in order to transfer the segmented contours into 2-D trajectories of discrete contour points. This enables the formulation of an adequate objective function and gives a detailed description of PE vibrations which is essential to reveal the dynamics of PE segment vibrations.

A variety of different algorithms that use point alignment to find an optimized transformation from a template image to a reference image have been suggested in the recent literature. However, most of these methods have difficulties if applied to PE oscillations. For a proper registration of PE contour deformations the following properties of PE oscillations are considered. Firstly, data points of the PE segment are located on a single closed contour, and they are consecutively numbered along this contour. Consequently, gradient information can be approximated and incorporated into the the optimization process. Secondly, the size and shape of the PE contours varies considerably during an oscillation cycle. Thus, one-to-one point assignments are in general not suitable to map contours from one frame to the next frame. This motivates the suggested advanced outlier handling as well as the introduction of 'dummy points' in regions where new parts of the contour in the reference frame have no appropriate counterpart in the template frame. Finally, information is used that a given PE contour appears again after a certain time interval. Thus, correlating PE con-

Figure 1.2: Oscillation cycles extracted from two high-speed movies (A,B) recorded during sustained phonation of a vowel. The fundamental frequencies of PE segment vibrations are 166 Hz and 125 Hz.
tours can be registered across an oscillation cycle which leads to a reduced error propagation in contrast to an exclusive frame-by-frame registration.

The paper is organized as follows: After a brief review of the relevant literature in the following section, an advanced contour matching approach for PE segment registration is described in Section 2.1. Results are discussed in Section 3, and the paper is concluded in Section 4 with a short summary of the results and an outlook on further research tasks.

## 1.1 Literature review

The following brief literature review concentrates on point based registration approaches that appear to be suitable for PE segment registration, and discusses their advantages as well as their shortcomings in the context of the application at hand. Since the suggested adaptive contour matching approach extends ideas from the robust point matching method of Chui and Rangarajan (2003), the corresponding results are reviewed in more detail.

The strategy using the obtained registration function as an approximation of the sought physical movement is already mentioned and used in Saadah et al. (1998) for the movement of the local folds in the larynx. The difference between our model and the model used in Saadah et al. (1998) is that we restrict the function space of the sought registration function, whereas in Saadah et al. (1998) a priori every two dimensional vector field \( u(x) \) is thought of as a possible movement \( f(x) = x + u(x) \). To ensure a certain degree of "smoothness", the angle between the displacement vectors of consecutive points on the contour is used as a regularization term. But with a movement described by a vector field no evaluation of the registration function outside of the considered points...
is possible. Therefore the gained information is very limited. We are going to present a method that overcomes these drawbacks because the solution obtained with this method is restricted to a “nice” function space.

This is achieved by an extension of the following registration algorithms. Landmark based registration algorithms like those suggested, for example, in Banerjee et al. (1995) use invariants of rigid transformations (such as angles or parallel lines) to find corresponding points (landmarks) in the reference and in the template image. Alternatively, landmarks can be selected manually (for example, by the physician) or automatically together with their correspondences. In medical image registration, morphologically characteristic landmarks can be used and automatically detected in both images. Consequently, landmark based algorithms are in most cases specialized to specific registration tasks, see Betke et al. (2003). Even though there are extensions to landmark based registration algorithms to non-rigid transformations like in Johnson and Christensen (2002) the problem of finding correspondences remains. Therefore landmark based registration algorithms are not suitable for our purpose. Moreover, it is generally impossible to decide how the data points are shifted between two consecutive frames of the video sequence since the movement of the soft tissue in the PE segment is to an extreme degree non-rigid.

The iterative closest point algorithm (ICP) first proposed in Besl and McKay (1992) uses a heuristic to find corresponding points. Each point in the first image is assigned to the nearest point in the second image, and the transformation is computed such that the sum of squared distances (SSD) between assigned points is minimal. This procedure is iterated until no further improvement of the SSD can be achieved. For rigid transformations, and given a good starting solution, this method works well and produces stable results. However, if non-rigid transformations are sought, or if a good initial guess is not available, the ICP generally converges to a stationary point or local minimum, which may in fact be far away from a reasonably good solution. There are various variants and improvements of the ICP, some of them are listed in Liu (2004).

In addition to the fact that the PE segment moves non-rigidly we have to address the problem of outlier points. The number of points on the contour may largely vary between consecutive frames because of the different contour length. This typically happens, for example, if the PE segment opens up in an area where it was previously closed and thus not visibly for the camera. In this situation a one-to-one point assignment is not meaningful. Note that this problem can not be overcome with a simple re-sampling step since the foregoing segmentation based on active contours (snakes) generates in general non-uniformly distributed points on the contour.

Chui and Rangarajan (2003) propose a robust point matching algorithm that incorporates a strategy for outlier handling in an ICP framework. The method is based on the formulation of a generalized assignment problem for finding corresponding points, and it involves the optimization of two types of variables: Discrete (binary) assignment variables and continuous transformation variables. Since our approach is closely related to the robust point matching algorithm, it will be reviewed in the following.
Problem 1.1 (Chui and Rangarajan (2003)). Let $X = \{x_1, \ldots, x_I\}$ and $Y = \{y_1, \ldots, y_J\}$ be sets of points in the template and the reference image, respectively, and let $\mathcal{F}$ be a finite dimensional function space containing candidates for the sought transformation $f$ between $X$ and $Y$. Then the robust point matching problem can be formulated as

$$\min \sum_{i=1}^{I} \sum_{j=1}^{J} z_{ij} \|y_j - f(x_i)\|^2 + \lambda \|Lf\|^2 - \xi \sum_{i=1}^{I} \sum_{j=1}^{J} z_{ij}$$

subject to

$$\sum_{i=1}^{I} z_{ij} = 1 \quad \forall j \in \{1, \ldots, J\}$$

$$\sum_{j=1}^{J} z_{ij} = 1 \quad \forall i \in \{1, \ldots, I\}$$

$$z_{ij} \in \{0, 1\} \quad \forall i \in \{1, \ldots, I+1\} \land j \in \{1, \ldots, J+1\}$$

$$f \in \mathcal{F},$$

where $\lambda, \xi \geq 0$ are fixed penalty parameters and $\|Lf\|^2$ is a measure for the smoothness of the transformation function $f \in \mathcal{F}$.

In formulation (1.1), the assignment variables $z_{ij}, i \in \{1, \ldots, I\}, j \in \{1, \ldots, J\}$, are equal to 1 if the point $x_i$ is assigned to the point $y_j$, and 0 otherwise. Thus the product $z_{ij} \|y_j - f(x_i)\|^2$ in the objective function implies the minimization of the distance $\|y_j - f(x_i)\|^2$ only for those pairs of points $y_j$ and $f(x_i)$ that are assigned to each other. If $z_{i,J+1} = 1$ for some $i \in \{1, \ldots, I\}$ or if $z_{I+1,j} = 1$ for some $j \in \{1, \ldots, J\}$, the point $x_i$ or $y_j$, respectively, is considered as an outlier and is excluded from the distance minimization. To avoid solutions with too many outliers, the term $-\xi \sum_{i=1}^{I} \sum_{j=1}^{J} z_{ij}$ is included in the objective function, where $\xi \in \mathbb{R}_+$ is a steering parameter that can be adapted with respect to the specific problem at hand. The smaller $\xi$ becomes the more points are identified as outliers, and in the case of $\xi = 0$ it is always optimal to consider every point as an outlier.

The function space $\mathcal{F}$ in (1.1) may be based on thin plate splines or radial basis functions, or it may be restricted to affine or rigid mappings. It should be noted that even if improvements of the ICP are applied to problem (1.1), it is very difficult or in some cases even impossible to find a global or even local optimum of the robust point matching problem for general non-rigid transformations.

But also for 'simple' function spaces $\mathcal{F}$, solving the robust point matching problem (1.1) remains computationally expensive. Due to the combination of integer assignment and continuous function interpolation variables, it belongs to the class of mixed integer programming problems (MIP), for which no polynomial time algorithm is known in general.

To overcome this difficulty, Chui and Rangarajan (2003) suggest a heuristic solution approach similar to the ICP, which solves the problem iteratively in the discrete and in the continuous variables. The resulting subproblems are an integer linear programming problem (in the assignment variables $z_{ij}$) and a convex optimization problem (in the transformation function $f$).

Note that due to the structure of the constraints the assignment matrices are contained in a polytope whose edges are binary matrices, as also noted in
Rangarajan et al. (1997a). The assignment subproblems are therefore totally unimodular, which implies that the assignment subproblem of (1.1) always has an integer optimal solution, even if the integrality constraints are relaxed to $z_{ij} \in [0,1] \forall i \in \{1, \ldots, I+1\}, j \in \{1, \ldots, J+1\}$.

The main drawback of this approach if applied to PE segment registration can be seen in the fact, that only the position of the points is taken into account, and no information about the shape of the contour is used. Numerical experiments showed that this often leads to wrong correspondences in the video frames, especially when the contour of the PE segment closes and opens again. In most of these cases we obtained assignments of points that are located on opposite sides of the gap. Due to these wrong assignments the corresponding transformations $f$ may be strange functions with (locally) very large gradients.

Another model formulation for the point matching problem is proposed by Zheng and Doermann (2006). An assignment is there chosen such, that neighboring structures are preserved and the mapping is “locally rigid”. But this may lead to arbitrary mappings and the movement of the PE Segment is not “locally rigid”.

This motivates the formulation of an alternative, adaptive contour matching problem and -algorithm that is capable of separating opposite contour lines, and that effectively uses all the information available in PE segment registration.

2 Methods

2.1 Registration of consecutive PE contours

While motivated by - and based on - the robust point matching algorithm of Chui and Rangarajan (2003), the adaptive contour matching approach described below modifies and extends this method in various ways to better incorporate the problem structure and the characteristics of PE segment registration. An advanced outlier handling, allowing for asymmetric assignments and dummy point introduction, is combined with an enhanced problem description incorporating gradient information. Particularly, introducing knowledge about the temporal properties of PE contour deformations facilitates proper registration results. The proposed registration procedure is also well suited for further registration problems that are based on a preceding segmentation step which yields finite and ordered point sets along closed contours.

2.1.1 Including normal information

One of the most problematic situations in PE segment registration is the registration of contours with incisions. When the distance between opposite curve segments of the contour is small, the robust point matching algorithm of Chui and Rangarajan (2003) frequently generates assignments across the shape (more or less independently of the chosen parameter settings). Measuring the registration error exclusively based on squared Euclidean distances between corresponding points is therefore not sufficient in this case. Therefore, we suggest
to use additionally a different quality measure, that is easily computable, that allows the separation of the inside from the outside of the contour, and that provides information about the local shape of the contour. These requirements are met, for example, by the angle between the normal vectors $n(y_j)$ to the contour, which can be easily approximated at the individual data points $y = y_j, j = \{1, \ldots, J\}$ using the two adjacent data points $y_{j-1}$ and $y_{j+1}$ that are immediately available due to the fixed ordering of points along the contour, see Figure 2.1.1. (To simplify notation, we set $y_{j+1} = y_1$ and $x_{i+1} = x_1$.)

![Figure 2.1: Approximation of the normal vector $n(y_j)$ and of the normalized tangent vector $\frac{t(y_j)}{\|t(y_j)\|}$ at $y_j$](image)

The additional integration of normal information, as also mentioned in Fry (1997), leads to a model formulation that considers both the local topology of the shape and the Euclidean distance between corresponding points. Unlike Fry (1997) who uses normal information as a distance measure for shape recognition, we want to augment the registration process with this concept.

Since the angle $\alpha$ between two normal vectors $n(y_j), n(f(x_i))$ is equal to the angle between the corresponding tangent vectors $t(y_j), t(f(x_i))$, the normal-angle-distance between two points $f(x_i)$ and $y_j$, $i = \{1, \ldots, I\}, j = \{1, \ldots, J\}$ can be approximated as follows:

$$d_{ij}^N = 1 - \cos \alpha = 1 - \frac{(f(x_{i+1}) - f(x_{i-1}))^T (y_{j+1} - y_{j-1})}{\|f(x_{i+1}) - f(x_{i-1})\| \|y_{j+1} - y_{j-1}\|}. \quad (2.1)$$

The relative influence of the normal distance as compared to the influence of the Euclidean distance in the objective function is controlled via a weighting
parameter $\delta \in [0, 1]$. This yields the following modified objective function:

$$
\min \delta \sum_{i=1}^{I} \sum_{j=1}^{J} z_{ij} d_{ij} + (1 - \delta) \sum_{i=1}^{I} \sum_{j=1}^{J} z_{ij} d_{ij} - \xi \sum_{i=1}^{I} \sum_{j=1}^{J} z_{ij},
$$

(2.2)

where $d_{ij} = \|f(x_i) - y_j\|^2$ is the Euclidean distance between the points $y_j$ and $f(x_i)$, $i = \{1, \ldots, I\}, j = \{1, \ldots, J\}$. This formulation has an interesting interpretation from the point of view of multicriteria optimization, where two or more objectives are optimized simultaneously. For all values of $\delta \in [0, 1]$, objective (2.2) can be seen as a weighted sums scalarization combining the two individual criteria $\sum_{i=1}^{I} \sum_{j=1}^{J} z_{ij} d_{ij}$ and $\sum_{i=1}^{I} \sum_{j=1}^{J} z_{ij} N_{ij}$ into one joint objective function. Similarly, the term $\sum_{i=1}^{I} \sum_{j=1}^{J} z_{ij}$ can be interpreted as a third objective aiming at the rejection of solutions with too many outliers, and weighted with the parameter $-\xi$. The resulting optimization problems are well posed for all feasible choices of $\delta$ and $\xi$, and they all yield a feasible registration. However, the registration result does for some parameter settings meet the requirements of the application in PE segment registration more than for others, even though we cannot define a 'best choice' for the parameters $\delta$ and $\xi$ from a theoretical point of view. In the computational tests a parameter $\delta = 0.85$ provided good results. Since the effect of the parameter $\xi$ is to a high degree dependent of the distance $l$ between consecutive points on the contour, we selected $\xi$ relative to $l$.

### 2.1.2 Asymmetric assignments: An alternative outlier handling

Different effects can corrupt the image quality during image acquisition, and wrong point detection may be a consequence of the segmentation procedure. To avoid that wrongly segmented points disturb the registration result, they should be identified as outliers and hence not taken into account when computing the transformation function. In this context, several authors suggested preprocessing steps using statistical methods for outlier rejection, see, for example, Besl and McKay (1992) and Wells III (1997), or special metrics that restrict the influence of outliers, Mount et al. (1998). Including a term for outlier handling directly in the assignment problem was suggested by Rangarajan et al. (1997b), and in a similar form by Zhang (1994) for free-form surfaces. This approach provides a comparably robust and elegant method that achieves good registration results even in the case of data or segmentation errors.

In PE segment registration we have a different situation. Due to the relatively clear structure of the PE segment with clearly visible boundaries in the video frames, the segmentation yields in general good results and there is no need to consider outliers caused by segmentation errors. However, in this case we are confronted with a different kind of 'outliers'. The contour of the PE segment may be – and is in practice – often of varying length. Therefore the number of detected points largely varies from frame to frame. Since between two frames
with a different number of points no one-to-one correspondence exists, a new kind of outlier handling has to be developed. Unlike the outliers caused by data errors, these ‘outliers’ should not be rejected during the registration process, but used instead as additional information defining the contour. This motivates the replacement of the classical assignment constraints in formulation (1.1) by the following constraints for an asymmetric and ambiguous assignment

\[
\sum_{i=1}^{I} z_{ij} \leq B_1 \\
\sum_{j=1}^{J} z_{ij} \leq B_2
\]  

(2.3)

with problem dependent and adaptable right hand sides \(B_1, B_2 \geq 1\). We suggest to select \(B_1\) and \(B_2\) depending on the ratio \(\frac{I}{J}\) between the number of points in the two frames that have to be registered. If \(I\) is significantly larger then \(J\) (for example, if \(I > 1.1 \cdot J\)), each data point \(y_j, j = \{1, \ldots, J\}\) may principally be assigned to more than one data point \(x_i, i = \{1, \ldots, I\}\). This can be realized, for example, by setting \(B_1 = 2\) and \(B_2 = 1\). However, in order to get more flexibility in the optimization process and thereby the ability to compensate for the fact, that the points are in general not uniformly distributed along the contour, we set \(B_1 = B_2 = 2\) whenever \(\frac{1}{1.1} \leq \frac{I}{J} \leq 1.1\), i.e.,

\[
B_1 = \begin{cases} 
1 & \text{if } J > 1.1 \cdot I \\
2 & \text{otherwise}
\end{cases} \quad B_2 = \begin{cases} 
1 & \text{if } I > 1.1 \cdot J \\
2 & \text{otherwise}
\end{cases}
\]

By using inequality constraints in (2.3) (instead of equality constraints as in the original assignment constraints, c.f. (1.1)), we can omit the outlier component in the assignment matrices (i.e., the \((I+1)\)-st and \((J+1)\)-st row and column, respectively). Nevertheless, the penalty term or, to be more precise, the ‘gratification term’ \(-\xi \sum_{i=1}^{I} \sum_{j=1}^{J} z_{ij}\) (c.f. (2.2)) is still needed in the objective function since otherwise the ‘empty’ assignment with \(z_{ij} = 0\) \(\forall i \in \{1, \ldots, I\}, j \in \{1, \ldots, J\}\) would always be an (unwanted) optimal solution to the problem.

As in the robust point matching problem (1.1), the number of outliers can be controlled by the parameter \(\xi\). Its effect depends, among others, on the distance between pairs of adjacent points in each data set (Euclidean plus normal-angle-distance). Since, due to the relative benefit of assignment and non-assignment in the objective function, only pairs of points with a combined distance smaller than \(\xi\) are aligned in an optimal solution, we select the parameter \(\xi\) depending on the magnitude of these point to point distances.

The discussion above leads to the following overall model for adaptive con-
tour matching:

\[
\begin{align*}
\min & \quad \delta \sum_{i=1}^{I} \sum_{j=1}^{J} z_{ij} d_{ij} + (1 - \delta) \sum_{i=1}^{I} \sum_{j=1}^{J} z_{ij} d_{ij}^N - \xi \sum_{i=1}^{I} \sum_{j=1}^{J} z_{ij} \\
\text{s.t.} & \quad d_{ij} = \|f(x_i) - y_j\|^2 \\
& \quad d_{ij}^N = 1 - \frac{(f(x_{i+1}) - f(x_{i-1}))^T (y_{j+1} - y_{j-1})}{\|f(x_{i+1}) - f(x_{i-1})\| \|y_{j+1} - y_{j-1}\|} \\
& \quad \sum_{i=1}^{I} z_{ij} \leq B_1 \quad \forall j \in \{1, \ldots, J\} \\
& \quad \sum_{j=1}^{J} z_{ij} \leq B_2 \quad \forall i \in \{1, \ldots, I\} \\
& \quad z_{ij} \in [0, 1] \quad \forall i \in \{1, \ldots, I\} \land j \in \{1, \ldots, J\}.
\end{align*}
\]

Registration results based on formulation (2.4) for different outlier handling parameters \( \xi \) are depicted in Figure 2.1.2.

2.1.3 Insertion of dummy points

As illustrated by the example in Figure 2.1.2, enforcing a complete registration with very few outliers may produce unrealistic transformations if the two considered frames are rather different. From a theoretical point of view, the registration shown in Figure 2.1.2(c) may nevertheless be optimal: It contains only very few outliers, and the distances between corresponding points are small. On the other hand, one of the requirements for the transformation was, that it should model the movement of the soft tissue in the PE segment. The function shown in Figure 2.1.2(c) clearly does not meet this requirement. In this example, the points in the two data sets do not represent the same morphological situation. In other words, the points, that are visible in the first frame, have not moved to cover all points in the second frame, and thus not according to the transformation shown in Figure 2.1.2(c). In fact, another gap has opened up in the PE segment that is visible in the second frame, but that was not visible in the first frame. The points on the contour of the new gap are in this sense 'new' points that have no correspondence in the first frame. Consequently, the movement of the left part of the segment that is present in both frames is better represented by the transformation function found with a smaller value of \( \xi \) as illustrated in Figure 2.1.2(b).

To obtain a realistic registration function that also simulates the movement of points in the case that a new gap opens up in the PE segment, we suggest the introduction of dummy points. Dummy points are artificial points that are included as correspondences for the new points in a new part of the segment. Allowing for 'pseudo assignments' (i.e., assignments including dummy points), information about the movement of the soft tissue at newly opened gaps can be included in the model.

A first step towards this goal is to formulate criteria for the detection of new gaps between two consecutive frames. This is realized by restricting the total number of consecutive outliers, and evaluating the obtained registration quality. If more than \( k \) consecutive points on the contour (with an appropriate
Figure 2.2: Registration results based on formulation (2.4) for different outlier parameters $\xi$ (frame 4)

parameter $k > 0$) are selected as outliers in an optimal solution of problem (2.4), we can assume that a new gap in the contour has opened up. The parameter $k$ should be selected depending on the number of points in each frame, $I$ and $J$. For our data sets with 80 to 300 data points a choice of $k \in [10, 15]$ led to the best results.

In the following discussion we assume that the points $y_{j}, \ldots, y_{j+\kappa}$ are outliers with some constant $\kappa > k$ and that a good pre-registration is known (e.g., from a previous registration step with a small value of $\xi$ or, as sufficient in the case of PE segment registration, given by the identity mapping).

Along the line starting from the centroid $S$ of the $\kappa$ outlier points in $Y$ and ending at the (closest) point $\bar{x} = \min_i ||S - x_i||$, we insert $\kappa$ dummy points $\tilde{x}_1, \ldots, \tilde{x}_\kappa$.

In every step two dummy points are placed, one of them shifted by a distance
of $\varepsilon$ to the left side of the line, the other to the right side. This is done in order to allow for registration functions that separate the dummy points and align them to opposite sides of the new contour. Using this approach, the tissue movement can be simulated very well by the obtained registration function.

2.1.4 Alternate convex search algorithm

Problem (2.4) is strongly non-convex and has in general a very large number of local minima, some of which correspond to very unrealistic registration functions (c.f. Figure 2.1.2(c) for an example). The determination of a global optimum with, for example, Branch and Bound methods is however in general very time-consuming. Since each video sequence of the PE segment consists of several hundreds of images that need to be registered, and since in this application the identity mapping is a very good starting solution, we have instead implemented a computationally efficient heuristic solution method based on 'Alternate Convex Search' (ACS). The method alternates between the determination of an optimal assignment for a given (and fixed) transformation function $f$, and the
consecutive optimization of $f$ for the fixed assignment determined before. This approach is iterated until no further improvements are made, or until some other stopping condition is satisfied. Note that both subproblems (the optimization of $f$ for given $z$, and the optimization of $z$ for given $f$) are convex and can thus be solved using methods of convex (or even linear) optimization. For the function space $\mathcal{F}$, from which the transformation function $f$ has to be selected, we have used thin-plate-splines as in Chui and Rangarajan (2003). The regularization term $\|Lf\|^2$ is based on the gradients of the transformation functions $f$. For a review of thin-plate splines and radial basis functions in image warping see Modersitzki (2004), Rohr et al. (2001) or Arad and Reisfeld (1995). Summarizing the discussion above, we can formulate an 'Adaptive Contour Matching' algorithm as follows:

Algorithm 2.1 (Adaptive Contour Matching).

1. Set $\lambda = \lambda^0 > 0$, $\xi > 0$, $k \in \mathbb{N}_+$, $0 < \tau < 1$, $0 \leq \delta \leq 1$

2. Optimize $z_{ij}$ with respect to

$$
\min \delta \sum_{i=1}^{I} \sum_{j=1}^{J} z_{ij}d_{ij} + (1 - \delta) \sum_{i=1}^{I} \sum_{j=1}^{J} z_{ij}d_{ij}^N - \xi \sum_{i=1}^{I} \sum_{j=1}^{J} z_{ij}$$

s.t. $d_{ij} = \|f(x_i) - y_j\|^2$

$$d_{ij}^N = 1 - \frac{(f(x_{i+1}) - f(x_{i-1}))^T(y_{j+1} - y_{j-1})}{\|f(x_{i+1}) - f(x_{i-1})\| \|y_{j+1} - y_{j-1}\|}$$

$$\sum_{i=1}^{I} z_{ij} \leq B_1 \quad \forall j \in \{1, \ldots, J\}$$

$$\sum_{j=1}^{J} z_{ij} \leq B_2 \quad \forall i \in \{1, \ldots, I\}$$

$$z_{ij} \in [0,1] \quad \forall i \in \{1, \ldots, I\} \land j \in \{1, \ldots, J\}$$

3. Optimize $f$ with respect to

$$
\min \sum_{i=1}^{I} \sum_{j=1}^{J} z_{ij} \|f(x_i) - y_j\|^2 + \lambda \|Lf\|^2
$$

s.t. $f \in \mathcal{F}$

4. WHILE no termination condition is satisfied, set $\lambda = \tau \lambda$, GOTO (2)

5. IF $\kappa > k$ consecutive points $y_j, \ldots, y_{j+k}$ are identified as outliers, GOTO (6), ELSE STOP

6. Determine the centroid $S$ of the $\kappa$ consecutive outlier points

7. Determine $\bar{x} = \text{argmin}_{x \in X} \|S - x\|

8. Insert $\kappa$ points $\bar{x}_1, \ldots, \bar{x}_\kappa$ along the line $(S, \bar{x})$ as illustrated in Figure 2.1.3

9. Set $X = X \cup \{\bar{x}_1, \ldots, \bar{x}_\kappa\}$ and reorder the data points

10. GOTO (2)
2.2 Registration of periodic contour deformations

Processing an entire high-speed sequence demands the registration of several hundreds of consecutive PE contours. A simple frame by frame registration would cause serious error propagation which would finally lead to a complete registration failure. However, PE contour deformations are not chaotic but follow a quasi-periodic vibration pattern. Particularly, consecutive oscillation cycles show strong geometric similarities of the deforming PE contours. Thus, a certain PE geometry reoccurs after a certain time interval which depends on the fundamental frequency of PE vibrations. Registration of PE contour deformations takes advantage of the quasi-periodic PE vibration pattern. Initially, for a high-speed sequence the fundamental frequency and cycle length of PE vibrations is determined. Following, the entire sequence is decomposed into consecutive cycles. Fig. 2.2 shows a PE oscillation cycle and schematically the further registration process. The beginning and ending of a cycle is defined by the maximum area (open-state) of the PE segment. The open-state possess a high amount of contour points which positively influences the performance of registration. Due to the knowledge that the last frame of a cycle resembles the first one, the two images are initially registered across the oscillation cycle. Following, starting from the first and last frame the remaining images of the PE cycle are forward and reversely registered, respectively.

The intermediate image of a PE cycle is registered twice. For this image, the analysis of differences between the forward and reverse registration enables an objective error estimation for PE registration.

\[
\text{err} = \frac{1}{K} \sum_{j=1}^{K} \| y_{f}^{j} - y_{r}^{j} \|_2, \tag{2.5}
\]

with \( K \) being the number of trajectories, \( y_{f}^{j} \) the endpoint of the \( j \)th trajectory in the intermediate frame after forward registration and \( y_{r}^{j} \) the endpoint in the intermediate frame after reverse registration. For all consecutive oscillation cycles the procedure is iterated which leads to a registration of an entire high-speed sequence.

2.3 Trajectories of discrete contour points

To visualize and validate the results of the obtained registration functions 2-D trajectories of discrete contour points can be computed. Therefore, a sequence of registrations over one oscillation cycle has to be computed. In the first frame we select points on the contour line of the PE segment as starting points for the tracing. After the computation of the registration function we apply it on the selected points. In order to avoid error propagation we project the mapped point orthogonally onto the contour line in the second frame. Since we select the considered oscillation cycle such that the PE segment is opened in the starting (and in the end) frame and (nearly) closed at a frame in the middle, the influence of errors can be diminished by changing the order of registrations. Instead of
Figure 2.5: Scheme for the registration of quasi-periodic PE contour deformations. Error propagation is reduced by an initial cross cycle registration, followed by successive forward/reverse registration steps. For the intermediate frame the comparison of the two registration results enables an error estimation of PE contour registration.

registering each frame to its consecutive frame, we determine the frame with the smallest open PE segment and register forward before and backwards after it. With this strategy the error made by merging the two parts of each trajectory is smaller than the error propagation made by simple forward registration, because the contour in the middle is smaller or even closed.

Algorithm 2.2 (Compute trajectories).

1. Choose an oscillation cycle, such that the PE segment is opened in the starting frame st\_frame and in the end frame end\_frame. Determine the frame m\_frame with the smallest opening.

2. Select the starting points x\_1, \ldots x\_n of the trajectories.

3. Register every frame from st\_frame to m\_frame \(-1\) with its consecutive frame according to algorithm 2.1.

4. Trace the selected points in st\_frame using the obtained registration functions.

5. Register st\_frame to end\_frame and compute the points corresponding to x\_1, \ldots x\_n in end\_frame.

6. Register every frame from end\_frame to m\_frame +1 with its prior frame according to Algorithm 2.1.

7. Trace the corresponding points in end\_frame using the obtained registration functions.
3 Results and Discussion

3.1 Registration of consecutive PE contours

All of the computations presented in this section have been run on an Intel Pentium 4, 2×3.2Ghz, 1000 MB RAM computer. The solution method was implemented in Matlab. All results were found by one single computation, using the identity mapping as starting solution for the transformation function $f$. Since the implemented ACS method generally converges to a local minimum in some attraction area of the starting solution, the results may vary if the computations are repeated with different starting solutions.

![Figure 3.1: Registration of frame 303 to frame 304 (parameter setting: $\xi = 0.21629, \delta = 0.85, \lambda^0 = 0.001, \tau = 0.1$)](image)

All problems considered are based on PE segment data obtained at the Department of Phoniatrics and Pediatric Audiology, University Hospital Erlangen, Germany. The registration shown in Figure 3.1 is in a sense an easy example of a registration problem. The structure and the morphology of the contour are nearly the same in the template and the reference image. The main difficulty in this example was the number of points: 291 in the template image and 268 in the reference image, respectively. Due to the relatively large number of points, the computation time of 121 seconds was one of the largest in our study.

In the example shown in Figure 3.2, the two opposite lines of the reference contour are in some areas relatively close to each other. Without including normal information, assignments across the contour are quite likely in this case. However, with the normal angle distance included in the objective function, the adaptive contour matching algorithm easily registers this closing movement of the PE segment.

The registration of frame 107 to 108 (shown in Figure 3.3) demonstrates again the necessity of including dummy points in the registration procedure. In this example, 64 dummy points were included. They were assigned to the new points on the right hand side of the contour. Since the registration process is restarted after the insertion of artificial points, the computation time is...
Figure 3.2: Registration of frame 261 to frame 262 (parameter setting: $\delta = 0.85$, $\lambda^0 = 0.001$, $\tau = 0.1$)

approximately twice the computation time without dummy points.

Figure 3.3: Registration of frame 107 to frame 108 (parameter setting: $\delta = 0.85$, $\lambda^0 = 0.001$, $\tau = 0.1$, 64 dummy points inserted)

The example shown in Figure 3.4 represents in some sense the inverse movement of the PE segment as the example shown in Figure 3.2. In this case the PE segment opens up, a movement that is much more difficult to represent by a registration function. Due to this opening of two very close contour lines, large gradients appear in the registration function. Since the regularization of the used thin-plate-splines is based on restricting their gradients, the parameter $\lambda^0$ has to be set to a very small value. The expanding registration function is illustrated in Figure 3.5.
Figure 3.4: Registration of frame 108 to frame 109 (parameter setting: $\delta = 0.85$, $\lambda^0 = 5e^{-10}$, $\tau = 0.1$)

Figure 3.5: Registration function for frame 108 to frame 109

3.2 Trajectories of discrete contour points

The presented registration procedure was successfully applied to high-speed recordings comprising about 350 consecutive frames. As the interpretation of abstract registration functions is not suitable for medical interpretation, the registered PE deformations are visualized using a set of discrete trajectories. Trajectories are derived from the computed registration functions and represent the 2-D deflection of a contour point with respect to time. Fig 3.6 shows a set of three trajectories which are displayed as black lines within the segmented PE contours during two oscillation cycles of a PE segment. The continuity of trajectories demonstrates the functionality of the suggested registration procedure.
It strictly follows the opening and closing of the pseudoglottal area.

Figure 3.6: 3-dimensional illustration of trajectories $T_1 - T_3$ of three contour points during a sequence length of 48 frames.

However, for medical purpose the interpretation of the temporal characteristics of the trajectories is hardly practicable within the 3-D visualization. In order to realize a proper assessment of oscillating PE contour deformations, a 2-D representation is suggested as shown in Fig. 3.7. For a single oscillation cycle the PE contours are displayed exclusively for the two open states, at the beginning and the end of the cycle. To capture the PE dynamics the lateral deflections of selected contour points are displayed as two-dimensional curves, which establish a connection between the two PE contours. The geometric pattern of the 2-D trajectories describe characteristically the PE vibrations at specified parts of the PE segment. Fig. 3.7(a) shows the 2-D representation of four PE trajectories for a single oscillation cycle derived from the registration results of the recording A. The trajectories start at four different points of the contour $c(frame = 1)$ and end at $c(frame = 20)$. The estimated registration error as defined in equation (2.5) is $0.7806 \pm 0.3444$ pixel. The progression of $T_1 - T_4$ is firstly inwardly directed and follows an elliptic path. The time dependency of the trajectories is exemplarily demonstrated within trajectory $T_3$ where the time axis is displayed. The highest deflections of PE tissue takes place at the dorsal left part of the PE segment. Fig. 3.7(b) shows the same PE trajectories obtained during a later oscillation cycle of the same high-speed recording. The principle pattern of the trajectories $T_1 - T_3$ is almost identical. The highest deflections occur again at the dorsal left part of the PE contour.
and all trajectories show an elliptic curvature. The stability of the trajectories reveal that the PE deformations do not show chaotic vibrations but follow a regular vibration pattern.

![Diagram of trajectories](image)

Figure 3.7: 2-dimensional illustration of four trajectories $T_1$-$T_4$ for two PE oscillation cycles registered from high-speed recording A. The fundamental frequency of PE segment oscillations is 200 Hz. (a) Cycle number = 4, (b) Cycle number = 10.

Fig. 3.8(a)-(b) visualize the registration results of the high-speed recording B. Again, a set of four characteristic trajectories are representatively selected to picture the vibration characteristics of the PE segment. In contrast to recording A, the shapes of the starting curve $c(frame = 1)$ and the end curve $c(frame = 28)$ differ considerably from each other. Nevertheless, a correct registration result could be achieved which is reflected by the progression of the trajectories $T_1$ - $T_4$ and the estimated registration error $0.9907 \pm 0.4216$ pixel. Here, the dominant deflections of PE oscillations equally takes place at the left part of the PE segment, reflected by $T_4$. The deflections of $T_1$ - $T_3$ are significantly reduced. The curves show a flat elliptic pattern and are firstly inwardly directed. During the PE closing they follow the same path as during the opening process and are thus almost located on top of each other. Again, the dominant vibration characteristics of the two oscillation cycles in sub-figure (a) and (b) strongly correlate to each other. The stability of the trajectories confirm the assumption, that during sustained phonation PE oscillations are reproducible and do not show chaotic behavior.

Up to the present, the analysis of PE dynamics simply focused on the description of the pseudoglottal area, either by visual inspection or image processing. The here presented registration procedure enables to derive further valuable information about the vibrating tissue of the PE segment. The visualization of the registration function by 2-D trajectories presents a new method to characterize PE dynamics. It permits to describe specifically the vibration pattern at different parts of the PE segment.

The time-dependent trajectory characteristics are affected by the morphology of the PE segment and depend consequently on the applied surgical tech-
Figure 3.8: 2-dimensional illustration of four trajectories $T_1-T_4$ for two PE oscillation cycles registered from high-speed recording B. The fundamental frequency of PE segment oscillations is 148 Hz. (a) Cycle number = 1, (b) Cycle number = 8.

4 Conclusions and future research

We have developed a robust registration method specifically tailored for PE segment registration. Our computational experiments show that 'Adaptive Contour Matching' can cope with PE segments of largely varying size and shape, and it integrates all information available in a combined objective function that can be adapted to specific data sets through an appropriate parameter selection. The approach is based on an advanced outlier handling strategy, including the option of ambiguous and asymmetric assignments, and the optimization process is driven by a combination of distance minimization, gradient alignment, and outlier rejection. An optimized registration function is found using an 'Alternate Convex Search' algorithm that alternates between the determination of optimal assignments for a given transformation function, and the optimization of the transformation function assuming a fixed assignment for corresponding points between the two frames being registered. Since the identity mapping is in general a good starting solution in the case of PE segment registration, the solution found with this approach models the movement of the soft tissue in the PE segment very well in all test data sets considered.

To extend this approach also to other registration tasks, and in particular to situations where no good starting solution is known, future research should include the development of a global optimization procedure for the adaptive contour matching problem. For this purpose, problem simplifications as, for example, a restriction to rigid transformation functions and a linearization of the objective function based on polyhedral distance measures may be used, and global solution strategies like Branch and Bound should be implemented.
Moreover, the selection of the weighting parameters $\delta$ and $\xi$ combining the three individual objective functions (measuring point distances, gradient alignment and outlier rejection, respectively) should be analyzed using, for example, concepts of multicriteria optimization. This allows a detailed analysis of the trade-off between the above mentioned criteria and of their impact on the found registration solution.

References


