Nasal Expiration Image Detection on Mirror Surface

Using Alternating Optimization

Kosuke Nakaharaa, Kazu Mishiba, Rieko Doi, Katsuya Kondo

aDept. Electrical Engineering and Computer Science, Tottori University, Tottori 680-8552, Japan
bDept. Medicine of Sensory and Motor Organs, Tottori University, Tottori 683-8504, Japan

*Corresponding Author: kondo@tottori-u.ac.jp

Abstract

ENT (ear, nose, and throat) doctors examine an infant nasal patency by using a nasal breath test. In nasal breath test, ENT doctors observe the fogging of mirror (the nasal expiration image) caused by the nasal expiration of the patient. Due to an immediate disappearance of the nasal expiration image, it is desirable to record the intensity of the nasal expiration image. The nasal expiration image is recorded by a video camera. The captured video is a mixture of video of the intensity of the nasal expiration image and video of background. The purpose of this paper is to separate video of the intensity of the nasal expiration image from captured video. The intensity of the nasal expiration image is estimated using an alternating optimization of nasal expiration image and background. Experimental results show that the proposed method can record the intensity of the nasal expiration image in a simulation environment.

Keywords: mirror image processing, nasal expiration image, alternating optimization.

1. Introduction

ENT (ear, nose, and throat) doctors examine an infant nasal patency by using a nasal breath test. In nasal breath test, ENT doctors observe the fogging of a Glatzel nasal mirror (the nasal expiration image) caused by the nasal expiration of the patient. Fig. 1 shows Glatzel nasal mirror. Due to an immediate disappearance of the nasal expiration image, it is desirable to record the data of the nasal expiration image. The data of the nasal expiration image denotes the density, the area, the direction, the shape, and others. The data of the nasal expiration image are used to quantitatively evaluate and to present the nasal breath test result. Therefore, a device that is capable of obtaining the data of the nasal expiration image has been invented. Haruna et al. proposed the liquid crystal rhinomanometer (LCR) (1). LCR utilizes a highly sensitive liquid crystal that changes color in response to the changes in temperature of the patient’s nasal expiration. Owing to the fact that the liquid crystal color shows the area, direction, and shape of the nasal expiration image, LCR can obtain these indices of the nasal expiration image. However, since the liquid crystal takes time to change color, it is difficult for LCR to obtain a real-time variation of data of nasal expiration image (2). To obtain real-time variation of the nasal expiration image data, Yasuhara et al. proposed a photo rhinometer (PRM) (3). PRM obtains the variation in the moisture deposit of nasal expiration. The moisture deposit of nasal expiration is generated on the surface of the glass by the patient’s nasal expiration. The variation in the moisture deposit of nasal expiration shows the area and speed variation of nasal expiration image in real time. The disadvantages of PRM are of large size and heavy weight. Therefore, Kondo et al. proposed a compact simple device for shooting the Glatzel nasal mirror from the left and right. This device uses image processing to obtain viewpoint from directly above of the nasal expiration image (4). However, this device does not obtain the nasal expiration image data. The density of the nasal expiration shows various data of the nasal expiration image and it is expressed as the intensity of the nasal expiration. Therefore, we want to obtain the

Fig. 1. Glatzel nasal mirror
intensity of the nasal expiration. The captured video is a mixture of video of the intensity of the nasal expiration image and video of background. Therefore, we need to separate video of the intensity of the nasal expiration image from captured video.

2. Modeling of Captured Video

In this paper, we assume four conditions. Four conditions make simplify the problem. First, a Glatzel nasal mirror is shot viewpoint from directly above. Second, the captured video is a grayscale movie because the color of nasal expiration is white when the color of ambient light is white. Third, the captured video is removed the lens distortion because camera parameters are known. Fourth, we are not considered some grooves of a Glatzel nasal mirror. In order to model the frame of the captured video, we assume that the captured video is a mixture of video of the intensity of the nasal expiration image and video of background. The nasal expiration image is generated by diffuse reflection. The diffuse reflection is caused by the water droplets adhering to the mirror surface. The color of diffuse reflection depends on the color of the ambient light. When the ambient light is white, the nasal expiration image is taken as white as the density increases. Therefore, the intensity of the nasal expiration image blends the white image with the background. Since doctors hold the video-capturing device in hand and record the nasal expiration image, the background moves. The background is assumed that intensity values do not vary temporally. Therefore, the frame of the captured video is defined as follows:

\[ l^t = α^t \circ F - (1 - α^t) \circ W(H^t)B, \tag{1} \]

where \( l \in \mathbb{R}^{h \times w} \) is the frame of the captured video, \( h \) is the height, \( w \) is the width, \( F \in \{1\}^{h \times w} \), \( B \in \mathbb{R}^{h \times w} \) is the background, \( α \in \mathbb{R}^{h \times w} \) is the intensity of the nasal expiration image in the range of 0–1, \( W \) is the transformation matrix, \( H \) is the homography matrix representing the motion of the background, \( t \) is frame number, and \( \circ \) is the Hadamard product. Our formulation method is similar to reflection removal \(^4\).

3. Proposed Method

3.1 Formulation

The unknown variables in Eq. (1) are \( H^t \), \( B \), and \( α^t \). We estimate the unknown \( H^t \) by using the M-estimator Sample Consensus (MSAC) algorithm \(^6\) and we estimate unknowns \( B \) and \( α^t \) by optimization method. First, we describe the estimate method for the homography matrix \( H^t \). The feature point with the SURF feature amount\(^b\) is detected from the captured video \( l^t(t = 1, …, n) \) where \( n \) is the number of frames. The feature points of the frame \( l^t \) correspond to the feature point of the frame \( l^t(t = 2, \ldots, n) \). Using the corresponded feature points and the MSAC algorithm, we estimate the homography matrices \( H^t(t = 2, \ldots, n) \)

Next, we describe how we estimate \( B \) and \( α^t \). We estimate them using the energy minimization framework. The energy minimization equation is defined as follows:

\[ \arg\min_{B, α^t} \{ E_D(B, α^t) + E_B(B) + E_α(α^t) \}, \tag{2} \]

where \( E_D(B, α^t) \) is a data term, \( E_B(B) \) is a term for the background \( B \), and \( E_α(α^t) \) is a term for the intensity \( α^t \) of the nasal expiration image. First, the data term \( E_D(B, α^t) \) is described as follows. We define the data term according to Eq. (1) as follows:

\[ E_D(B, α^t) = \sum_{t=1}^{n} \| l^t - α^t \circ F - (1 - α^t) \circ W(H^t)B \|_{F}^2, \tag{3} \]

where \( \| \cdot \|_F \) is frobenius norm. Next, the term \( E_B(B) \) for the background \( B \) is described as follows. The background \( B \) is an image containing noise and the intensity of \( B \) is constrained to be in the range of 0–1. We need to reduce the noise effect. Therefore, the term \( E_B(B) \) for the background \( B \) is defined, as follows:

\[ E_B(B) = γ_1 \| \nabla \ast B \|_{F}^2 \]

subject to \( 0 \leq B(i, j) \leq 1, \tag{4} \]

where \( \nabla = \begin{bmatrix} 1 & 1 & 1 \\ 1 & -2 & 1 \\ 1 & 1 & 1 \end{bmatrix} \ast \) is convolution operator, \( i \) is the \( x \)-coordinate of the image, \( j \) is the \( y \)-coordinate of the image, and \( γ_1 > 0 \) is the weight. Finally, the term \( E_α(α^t) \) for the intensity \( α^t \) of the nasal expiration image is described as follows. Since the density of the fogging of mirror changes continuously both spatially and temporally, the intensity \( α^t \) of the nasal expiration image is both spatially and temporally smooth. Therefore, the term \( E_α(α^t) \) for the intensity \( α^t \) of the nasal expiration image is defined as follows:

\[ E_α(α^t) = γ_2 \sum_{t=1}^{n} \| \nabla \ast α^t \|_{F}^2 + γ_3 \sum_{t=2}^{n} \| α^t - α^{t-1} \|_{F}^2, \tag{5} \]

subject to \( 0 \leq α^t(i, j) \leq 1, \)

where \( γ_2 \) and \( γ_3 > 0 \) are the weights.
3.2 Optimization

Equation (2) is solved using alternating optimization. The intensity $\alpha^t$ of the nasal expiration image of the frame $I^t$ is assumed to be 0. The initial values are described as follows. The initial values $B_0 \in \mathbb{R}^{h \times w}$ of the background is defined as the frame $I^0$ and $a_0^t$ is obtained as follows:

$$a_0^t = (I^t - W(H^t)B_0) \cdot (F - W(H^t)B_0),$$

(6)

where $\cdot$ is elementwise division. We describe the optimization as follows. First, we fix $\alpha^t$ and optimize $B$. Using Eqs. (3) and (4), the optimization formula for $B$ is expressed as follows:

$$\arg\min_B \left\{ \sum_{t=1}^n \| I^t - a^t \ast F - (1 - a^t) \ast W(H^t)B \|_F^2 + \gamma_1 \| \nabla B \|_F^2 \right\}$$

(7)

subject to $0 \leq B(i,j) \leq 1$.

Equation (7) is solved using metric projection with the BOX constraint. Next, we fix $B$ and optimize $\alpha^t$. Using Eqs. (3) and (5), the optimization formula for $\alpha^t$ is expressed as follows:

$$\arg\min_{\alpha^t} \left\{ \sum_{t=1}^n \| I^t - a^t \ast F - (1 - a^t) \ast W(H^t)B \|_F^2 + \gamma_2 \sum_{t=1}^n \| \nabla \ast a^t \|_F^2 + \gamma_3 \sum_{t=1}^n \| a^t - a^{t-1} \|_F^2 \right\}$$

(8)

subject to $0 \leq a^t(i,j) \leq 1$.

Similar to the optimization formula for $B$, Eq. (8) is solved using metric projection with the BOX constraint. The iteration process of alternating optimization ends when the set number of updates elapses. When the background is transformed by the homography matrix, the border of the image has no value. Therefore we can not estimate the end of the intensity of the nasal expiration image. The nasal expiration image is not generated at the border of the image. Therefore, the border of the image is excluded from the optimization processing.

4. Experiment

To evaluate our method, we performed experiment with simulation data. The purpose of this experiment is to verify two effectiveness. First effectiveness is robustness against the background motion. Second effectiveness is robustness of our optimization formula against noise of the captured

video. First effectiveness is verified by comparing the result of Eq. (6) with the result of the formula that does not consider the background motion, which is expressed as follows:

$$a_{dit}^t = (I^t - B_0) \cdot (F - B_0).$$

(9)

Eq. (9) is a formula of removing $W(H^t)$ from Eq. (6). Second effectiveness is verified by comparing the result of optimization with the result of Eq. (6). Since there is no state-of-the-art method of recording the intensity of the nasal expiration image, we compared the result of optimization with the result of Eq. (6).

4.1 Simulation Experiment

To evaluate how close the intensity of the nasal expiration image is to the ground truth, we created simulation data. The purpose observing the variation in the intensity of the nasal expiration using captured nasal expiration images is to simulate the intensity of the nasal
expiration. Fig. 2 shows the image obtained by taking a nasal expiration image viewpoint from directly above. The background is uniform. The nasal expiration image is often elliptical, as shown in Fig. 2 (1). Therefore, using the sigmoid function and the equation of an ellipse, the variation in the intensity of the nasal expiration image in the x-axis is defined as follows:

\[
z_x(x, y) = \begin{cases} 
1 & -c_x \leq x < 0, 0 \leq y \leq 2b_x \\
1 + \exp(-g(x + x_0)) & 0 \leq x \leq c_x, 0 \leq y \leq 2b_x \\
1 + \exp(g(x - x_0)) & 0, \\
0 & \text{otherwise}
\end{cases}
\] (10)

where \( z_x \) is the intensity of the nasal expiration image, \( g \) is the gain of the sigmoid function, \( a \) is the radius of the ellipse in the x-axis direction, \( b \) is the radius in the y-axis direction, and \( c_x \) is the coordinate of the end of nasal expiration image. The experiments assumed \( g = 0.2 \), \( a_x = 30 \), \( b_x = 60 \), and \( c_x = 60 \). Figure 3 shows a line profile of Eq. (10) when \( y = 60 \). The variation in the intensity of the nasal expiration image in the y-axis direction is high near the nose. Therefore, the variation of the intensity of the nasal expiration image in the y-axis direction is defined using the elliptic equation as follows:

\[
z_y(y) = \begin{cases} 
1 - \frac{y^2}{b_y^2} & 0 \leq y \leq b_y \\
0 & y > b_y
\end{cases}
\] (12)

where \( b_y \) is the radius of the ellipse in the y-axis direction. The experiments assumed \( b_y = 120 \). Figure 4 shows a line profile of Eq. (12). As shown in Fig. 2, the intensity of the nasal expiration image increases in the first few frames and then reach a plateau. Therefore, the temporal variation of the intensity of the nasal expiration image is defined as follows using the exponential function (exp):

\[
z(t) = 1 - \exp(-t/p),
\] (13)

where \( p \) is a constant. The experiments assumed \( p = 5 \) and a line profile of Eq. (13) is shown in Fig. 5. Using Eqs. (11), (12), and (13), we define the intensity of the nasal expiration image as follows:

\[
z = k_{xz}z_xz_t,
\] (14)

where \( k \) is a constant. The experiments assumed \( k = 0.5 \). The nasal expiration image exists in the center of the image and the orientation is oblique. Therefore, the intensity of the nasal expiration image \( z \) are rotated and translated. The background of the simulation data is obtained by capturing a ceiling. The homography matrix \( H \) was used to move the background. The homography matrix \( H \) is an arbitrary matrix. The simulation data is assumed that there is no nasal expiration image in the first frame. Example of the simulation data is shown in Fig. 6. In order to evaluate how close the intensity of the nasal expiration image is to the ground truth, we evaluate the results using MSE.
4.2 Results

In our implementation, the parameters in Eqs. (7) (8) were $\gamma_1 = 0.5$, $\gamma_2 = 0.1$ and $\gamma_3 = 0.3$. The experiment results are shown in Table 1 and Figs. 7, 8 and. Table 1 lists the MSE of the image cropped around the nasal expiration image and Figs 7, 8 show the images cropped around the nasal expiration images. On the contrary to results using Eqs. (6) and (9), our result exhibits higher performance both visually and quantitatively. Since the motion of the background is not considered, Eq. (9) exhibits a lower performance. Eq. (6) considers the motion of the background. However, Eq. (6) does not consider the prior of the background and the intensity of the nasal expiration image. The result with Eq. (6) exhibits a lower performance in comparison with our result. The proposed method produces a satisfactory result in most cases.

Table 1. Comparison of MSE of the nasal expiration image

<table>
<thead>
<tr>
<th></th>
<th>Eq. (9)</th>
<th>Eq. (6)</th>
<th>Proposed</th>
</tr>
</thead>
<tbody>
<tr>
<td>2nd frame</td>
<td>0.013</td>
<td>$4.03 \times 10^{-5}$</td>
<td>$7.39 \times 10^{-5}$</td>
</tr>
<tr>
<td>5th frame</td>
<td>0.008</td>
<td>$5.32 \times 10^{-5}$</td>
<td>$1.79 \times 10^{-5}$</td>
</tr>
<tr>
<td>15th frame</td>
<td>0.015</td>
<td>$32.7 \times 10^{-5}$</td>
<td>$2.85 \times 10^{-5}$</td>
</tr>
<tr>
<td>25th frame</td>
<td>0.007</td>
<td>$88.3 \times 10^{-5}$</td>
<td>$5.72 \times 10^{-5}$</td>
</tr>
<tr>
<td>30th frame</td>
<td>0.011</td>
<td>$2.05 \times 10^{-5}$</td>
<td>$2.77 \times 10^{-5}$</td>
</tr>
<tr>
<td>Average (2-31 frame)</td>
<td>0.010</td>
<td>$14.8 \times 10^{-5}$</td>
<td>$2.79 \times 10^{-5}$</td>
</tr>
</tbody>
</table>
5. Conclusions

We proposed a separate method of the intensity of the nasal expiration image from captured video by using alternating optimization of the intensity of nasal expiration image and background. Alternating optimization is solved using metric projection to BOX constraints. Simulation results show that the proposed method can record the intensity of the nasal expiration image.

References


Fig. 8. Comparison of the nasal expiration image (15th frame).