Contents lists available at ScienceDirect





CrossMark

Precision Engineering

journal homepage: www.elsevier.com/locate/precision

Registration of infrared transmission images using squared-loss mutual information

Tomoya Sakai^{a,*}, Masashi Sugiyama^a, Katsuichi Kitagawa^b, Kazuyoshi Suzuki^b

^a Tokyo Institute of Technology, 2-12-1 O-okayama, Meguro-ku, Tokyo 152-8552, Japan ^b Toray Engineering Co., Ltd., 1-1-45 Oe, Otsu, Shiga 520-2141, Japan

ARTICLE INFO

Article history: Received 30 April 2014 Received in revised form 18 August 2014 Accepted 31 August 2014 Available online 16 September 2014

Keywords: Infrared transmission image Image registration Squared-loss mutual information

ABSTRACT

Infrared light allows us to measure the inner structure of opaque samples such as a semi-conductor. In this paper, we propose a method of registering multiple infrared transmission images obtained from different layers of a sample for 3D reconstruction. Since an infrared transmission image obtained from one layer is contaminated with defocused images coming from other layers, registration with a standard similarity metric such as the squared error and the cross correlation does not perform well. To cope with this problem, we propose to use the *squared-loss mutual information* as an alternative similarity measure for registration, which is more robust against noise than ordinary mutual information. The practical usefulness of the proposed method is demonstrated in simulated and actual experiments.

© 2014 Elsevier Inc. All rights reserved.

1. Introduction

Non-destructive inspection of precision instruments such as semi-conductors is one of the most important manufacturing processes in precision industries, and measuring inner structure of opaque objects by *infrared light* is a promising means for this purpose. However, images obtained by infrared light from different layers of a sample often suffer misalignment because of asynchronous scanning of images from different layers. In this paper, we therefore consider the problem of registering multiple infrared transmission images obtained from different layers of a sample, and propose a new practical algorithm for 3D reconstruction. Our proposed method can be used, e.g., for identifying the position of defects in semi-conductor samples and thus can provide more precise information of the inner structure for inspection.

Related image registration problems have been explored, e.g., in photolithography processes for aligning circuit patterns and masks [1] and in pattern and photo-mask inspection for comparing target and reference images [2]. On the other hand, the image registration problem we are tackling in this paper is much more challenging than the previous works because an infrared transmission image obtained from one layer is contaminated with defocused images coming from other layers. For this reason, standard *linear* similarity

* Corresponding author. Tel.: +81 3 5734 2699.

E-mail addresses: sakai@sg.cs.titech.ac.jp (T. Sakai),

sugi@cs.titech.ac.jp (M. Sugiyama), BXM02060@nifty.ne.jp (K. Kitagawa), kazuyoshi_suzuki@toray-eng.co.jp (K. Suzuki).

http://dx.doi.org/10.1016/j.precisioneng.2014.09.001 0141-6359/© 2014 Elsevier Inc. All rights reserved. metrics such as the *sum of squared differences* (SSD) and the *normalized cross-correlation* (NCC) [3–8] are not suitable to registration of infrared transmission images.

Mutual information (MI) [9], which is a quantity of interest in the *information theory* community, allows us to capture non-linear variations between two random variables:

$$MI := \iint p(x, y) \log \frac{p(x, y)}{p(x)p(y)} dx dy,$$
(1)

where p(x, y) is the joint probability density of x and y, and p(x)and p(y) are the marginal probability densities of x and y, respectively. MI is always non-negative, and takes zero if and only if x and y are statistically independent. Thus, MI measures the dependency between x and y, which describes more precise "relation" than linear correlations. For example, if $y = x^2$, x and y are uncorrelated but they are still statistically dependent and thus MI takes a strictly positive value. Note that MI is reduced to a linear correlation measure when x and y follow the centered Gaussian distributions. Thus, MI can be regarded as a generalization of linear correlations to non-Gaussian random variables.

Given this superior detectability of statistical dependency, MI should be regarded as a suitable similarity measure for registration of infrared transmission images. However, because of the log function included in the definition of MI, which is an extremely steep function near the origin, MI is highly sensitive to noise and outliers [10].

To overcome the excessive sensitivity of MI, a variant of MI called *squared-loss mutual information* (SMI) [11] was introduced:

SMI :=
$$\frac{1}{2} \iint p(x)p(y) \left(\frac{p(x,y)}{p(x)p(y)} - 1\right)^2 dxdy.$$
 (2)

SMI is also always non-negative and takes zero if and only if *x* and *y* are statistically independent. Thus, SMI can be used as an alternative to MI for evaluating statistical dependency, without suffering the "log" problem. Furthermore, thanks to the simple squared-difference expression of SMI, it can be *analytically* approximated from samples in a statistically optimal way, and this analytic SMI approximator allows explicit computation of its derivative [12]. This is a highly useful property in image registration, because eventually we want to register images so that SMI is maximized with respect to some image transformation parameters. For these reasons, we propose to use SMI as our dependency measure for registration of infrared transmission images.

After reviewing an SMI approximator in Section 2, we describe our SMI-based image registration method in Section 3. Its performance is experimentally evaluated in Section 4, and we conclude in Section 5.

2. SMI approximation

Since SMI defined by Eq. (2) contains unknown probability densities p(x, y), p(x), and p(y), its value cannot be directly computed. In this section, we review how SMI is approximately computed from paired samples $\{(x_i, y_i)\}_{i=1}^n$ independently following p(x, y).

A naive way to approximately compute SMI is to separately estimate the densities p(x, y), p(x), and p(y) from samples $\{(x_i, y_i)\}_{i=1}^n$, and plug the estimated densities $\hat{p}(x, y)$, $\hat{p}(x)$, and $\hat{p}(y)$ in the definition of SMI. However, such a plug-in approach is known to perform poorly, because the first step of estimating densities is performed without regards to the second step of plugging them in SMI. More specifically, dividing $\hat{p}(x, y)$ by $\hat{p}(x)$ and $\hat{p}(y)$ can significantly magnify the estimation error of $\hat{p}(x, y)$ when $\hat{p}(x)$ and $\hat{p}(y)$ take a small value [13].

To cope with this problem, the direct SMI approximator called *least-squares mutual information* (LSMI) [11,10] was proposed. LSMI directly estimates the density ratio function,

$$r(x,y) := \frac{p(x,y)}{p(x)p(y)},\tag{3}$$

without individually estimating each density. More specifically, the density ratio r(x, y) is modeled by the following *multiplicative kernel* model [14]:

$$r_{\Theta}(x, y) := \sum_{i,j=1}^{n} \Theta_{i,j} K(x, x_i) L(y, y_j),$$
(4)

where K(x, x') and L(y, y') are kernel functions for x and y, and Θ is a parameter matrix whose (i, j)-element is $\Theta_{i,j}$. Below, we focus on the Gaussian kernel for K(x, x') and L(y, y'):

$$K(x, x') := \exp\left(-\frac{(x - x')^2}{2\sigma^2}\right),\tag{5}$$

$$L(y, y') := \exp\left(-\frac{(y-y')^2}{2\sigma^2}\right),\tag{6}$$

where σ denotes the Gaussian bandwidth. When the sample size n is too large, we may reduce the number of kernel bases in the model (4) by, e.g., randomly selecting a subset. In our experiments, we randomly choose 100 kernel bases.

The parameter $\boldsymbol{\Theta}$ is determined so that the following squared error *J* is minimized:

$$J(\boldsymbol{\Theta}) := \frac{1}{2} \iint p(x)p(y)(r_{\boldsymbol{\Theta}}(x,y) - r(x,y))^{2} dx dy$$
$$= \frac{1}{2} \iint p(x)p(y)r_{\boldsymbol{\Theta}}(x,y)^{2} dx dy - \iint p(x,y)r_{\boldsymbol{\Theta}}(x,y) dx dy$$
$$+ C, \tag{7}$$

where *C* is a constant independent of $\boldsymbol{\Theta}$. By ignoring the irrelevant constant *C*, approximating the expectations with the empirical averages, and including the ℓ_2 -regularizer for avoiding overfitting, the LSMI solution $\hat{\boldsymbol{\Theta}}$ was shown to satisfy the following equation [14]:

$$\frac{1}{n^2} \mathbf{K}^2 \widehat{\mathbf{\Theta}} \mathbf{L}^2 + \lambda \widehat{\mathbf{\Theta}} = \frac{1}{n} \mathbf{K} \mathbf{L},\tag{8}$$

where $K_{i,j} = K(x_i, x_j)$, $L_{i,j} = L(y_i, y_j)$, and $\lambda \ge 0$ denotes the regularization parameter that is determined by *cross-validation* with respect to *J* [11]. The above equation is called the *discrete-time Sylvester* equation [15], and is known to be solved in $\mathcal{O}(n^3)$ time, e.g., via eigendecomposition.

Finally, based on another expression of SMI,

$$SMI = \frac{1}{2} \iint p(x, y)r(x, y)dxdy - \frac{1}{2},$$
(9)

the LSMI estimator is given as

$$LSMI := \frac{1}{2n} tr(\mathbf{K}\widehat{\boldsymbol{\Theta}}\mathbf{L}) - \frac{1}{2}.$$
 (10)

3. Image registration with LSMI

Let us consider the problem of registering images X and Y: we transform image X to \widetilde{X} (for example, by rotation or translation), so that \widetilde{X} "matches" Y as much as possible. As the matching score, we use LSMI between \widetilde{X} and Y. We regard \widetilde{x} as a pixel value of image \widetilde{X} and y as a pixel value of image Y, and we generate paired samples $\{(\widetilde{x}_i, y_i)\}_{i=1}^n$ by coupling pixel values at corresponding points in the images.

More specifically, our goal is to maximize LSMI with respect to an image transformation parameter from X to \tilde{X} . Here, we use a gradient-based method, such as a gradient ascent method and a quasi-Newton method, to find a (local) maximizer of LSMI. The gradient of LSMI is given by

$$\nabla \text{LSMI} = \frac{1}{n^2 \sigma^2} \text{tr}(\widehat{\boldsymbol{\Theta}}^\top \widetilde{\boldsymbol{K}} \boldsymbol{Q} \widehat{\boldsymbol{\Theta}} \boldsymbol{L}^2) - \frac{1}{n \sigma^2} \text{tr}(\boldsymbol{Q} \widehat{\boldsymbol{\Theta}} \boldsymbol{L}), \qquad (11)$$

where

$$\widetilde{K}_{i,j} := K(\widetilde{x}_i, \widetilde{x}_j), \tag{12}$$

$$Q_{i,j} := \widetilde{K}_{i,j}(\widetilde{x}_i - \widetilde{x}_j) \nabla \widetilde{x}_i, \tag{13}$$

and $\nabla \tilde{x}_i$ denotes the gradient of the pixel value at \tilde{x}_i .

Below, for simplicity, we focus on translation as image transformation.¹ Then, the pixel values of the transformed image \widetilde{X} are given by

$$\widetilde{x}(u,v) := x(u - w_u, v - w_v), \tag{14}$$

¹ Note that our framework can handle any transformation as long as it is smooth with respect to transformation parameters.

Download English Version:

https://daneshyari.com/en/article/804500

Download Persian Version:

https://daneshyari.com/article/804500

Daneshyari.com