Optics Communications 283 (2010) 4840–4845

Contents lists available at ScienceDirect

Optics Communications

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

An optimized in vivo multiple-baseline stereo imaging system for skin wrinkles

Onseok Lee [a,b], Gunwoo Lee [a,b], Jangseok Oh [c], Mingi Kim [c,*], Chilhwan Oh [a,b,d,*]

[a] Biomedical Engineering, Biomedical Science of Brain Korea 21, Korea University College of Medicine, Anam-dong, Seongbuk-gu, Seoul, 136-705, Republic of Korea
[b] Research Institute for Skin Image, Korea University Medical Center, Guro2-dong, Guro-gu, Seoul, 152-703, Republic of Korea
[c] 3D Information Processing Laboratory, Korea University, Anam-dong, Seongbuk-gu, Seoul, 136-705, Republic of Korea
[d] Department of Dermatology, Korea University Guro Hospital, Korea University College of Medicine, Guro2-dong, Guro-gu, Seoul, 152-703, Republic of Korea

ARTICLE INFO

Article history:
Received 18 February 2010
Received in revised form 1 July 2010
Accepted 1 July 2010

Keywords:
Stereo image
Baseline
Skin
Disparity
Graph cut

ABSTRACT

Recent developments in computer vision have attempted to mimic human visual physiology. One application of this technology is the evaluation of skin wrinkles. We have developed a non-convergence stereo system that can be calibrated in vivo and can be controlled to the level of microns for baseline. We are able to obtain more accurate 3D information by calibrating nonlinear interrelations between the disparity of object and depth information.

1. Introduction

Stereo vision can be used to obtain three-dimensional (3D) information (disparity), which represents the differences in position between two points projected in 3D space [1,2]. The disparity d is related to the depth z by

$$d = BF\frac{1}{z} \tag{1}$$

where B and F are the baseline and focal lengths, respectively. In other words, we can obtain better precision in objects with a limited range of depth information because it can be expressed as larger disparity with the baseline, the distance between the optical axes of two cameras. However, the larger the baseline, the larger the stereo matching line, which can degrade correctness, leading to a tradeoff between precision and correctness [3]. Additionally, the above formula shows that depth information, expressed as a disparity, is a nonlinear relation. Thus, to obtain correct and quantitative 3D information using stereo images, both the correct choice of baseline and the calibration of the disparity are important [4].

Stereo camera systems are necessary to obtain 3D information from stereo images. These can be divided into non-convergence and convergence camera models [1,2,5]. Non-convergence cameras obtain stereo images without vertical range by arranging the optical axes of two parallel cameras [6,7]. As it is difficult to satisfy this condition, convergence camera models with vertical range are normally used. These cases require a complicated stereo matching procedure, and rectification is necessary to remove vertical range [1,2]. This has the effect of improving calculation speed and correctness by changing 2D problems with vertical and horizontal aspects to 1D problems by removing the vertical aspect.

Recently, much research has attempted to replicate the human visual system via computer [8]. This research has many applications, including evaluating skin diseases in the medical field [5,9–11]. However, existing systems cannot select suitable baseline images for calibration because the baseline is fixed. These techniques use a convergence camera model to create stereo images by using a mirror connected stereo system with multiply angles so that a weak impact can twist the reflection angle [12]. The technology allows the camera to shift along the X, Y, and Z axes delicately, allowing it to control the baseline and obtain non-convergence stereo images. However, it cannot obtain right and left images simultaneously, so it is improper for medical applications that need to obtain *in vivo* stereo images. Obtaining images by replicating living skin using silicon can remove the right/left image change in the region of interest (ROI) caused by breathing and other movements, but this brings about errors in 3D information caused by minute bubbles in the silicon that are created by manufacturing replicas.

Accordingly, obtaining more correct and quantitative 3D information using stereo images in the medical field requires: (1) a system available for minute baseline control at high powered images to

* Corresponding authors. Oh is to be contacted at Department of Dermatology, Korea University Guro Hospital, Korea University, College of Medicine, 97 Guro-dong gil, Guro-dong, Guro-gu, Seoul, 152-703, Republic of Korea. Tel.: +82 2 2626 1894; fax: +82 2 2626 1900. Kim, 3D Information Processing Lab., Korea University, 126-1 Anam-dong 5ga, Seongbuk-gu, Seoul, 136-701, Republic of Korea. Tel./fax: +82 2 3290 3977.

E-mail addresses: mgkim@korea.ac.kr (M. Kim), choh@korea.ac.kr (C. Oh).

0030-4018/$ – see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.optcom.2010.07.005

obtain a suitable baseline for a stereo image of living skin; (2) a system to obtain non-convergence stereo images with correct fast matching that does not require a rectification process; and (3) an *in vivo* system to obtain right and left images simultaneously without rectification.

We constructed a system that satisfies all of the above conditions and deduces correct 3D information by calibrating the nonlinear relationship between disparity and depth information.

2. Materials and methods

Fig. 1 is a schematic of the main body required to realize a system appropriate for the purposes of this study. The piece used to obtain the right and left image was made independently, and the zoom lens (low magnification Zoom 7000, NAVITA®, Japan) used to control magnification and the charge-coupled device (CCD) camera (GEV GP-3780C, GEViCAM Inc., USA) used to obtain images were made according to common specifications. The part to obtain the left images (hereafter known as the shifting part) is attached to a manual stage (Parker Hannifin-Daedal Division, USA) and was designed to shift from a minimum of 5.08 μm (0.0002 inch) to a maximum of tens of millimeters from side to side. This allows operators to obtain baseline values to obtain correct 3D information depending on the type of skin being visualized.

The piece to obtain right images (hereafter known as the fixed part) was designed to not shift. The shifting and fixed parts are arranged vertically to avoid a constitutional conflict which would prevent obtaining two images simultaneously, and beam splitter (Green Optics Co., Ltd., Korea) is installed at the crossing point of each optical line to detect the images. The merits of this design are that it improves the correctness and speed of stereo matching because one can obtain non-convergence stereo images. Additionally, *in vivo* images can also be obtained since you can capture stereo images simultaneously.

This main body of the apparatus can obtain *in vivo* images of living skin as it selects points for calibration using a shiftable stand (Articulating Arm Boom Stand, EO® Edmund Optics Inc., USA) in 3D space.

The images captured by the CCD camera are sent to a computer using a Dual Port Gigabit Ethernet Server NIC (INTEL® PRO/1000, USA)

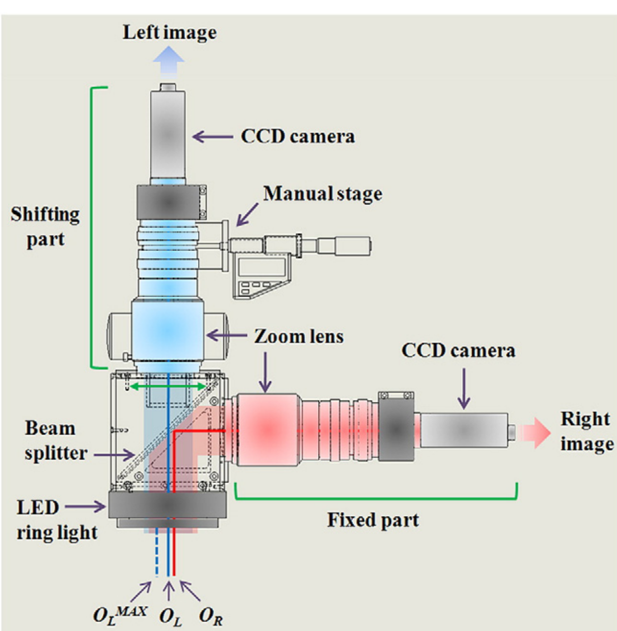

Fig. 1. Schematic for the main apparatus of the *in vivo* multiple-baseline stereo imaging system.

for speed. The main body uses a light emitting diode (LED) ring light and an LED power supply (SEOKWANG CO., LTD., Korea) to allow for minute focusing control. Additionally, we included a diffusion filter (Hongseong optics CO., LTD., Korea) to distribute the light evenly. In this system, operators should not expect even light distribution because there is a difference in the intensity of each image caused by refraction and reflection due to the fact that the beam splitter and light guide are round. In the meantime, the stereo matching point is a main factor in deciding a pixel's intensity in the right and left images [2]. Uneven light distribution leads to decreased correctness in stereo matching, and so the image is calibrated by subtraction from a white reference image.

Stereo images from this system show a disparity map converted to intensity from the disparity of matching point through stereo matching. Disparity can be calibrated in the process of researching the right image based on left image because the similarity between corresponding points of the two images should be judged in advance. We obtained the disparity map through a graph cut that is known to give excellent results [13–15].

Our results show that we can select a baseline after calculating the disparity map based on *in vivo* multiple-baseline stereo images of skin wrinkles using a non-convergence camera system and comparing the results from each baseline.

As mentioned above, the relationship between disparity and depth information is nonlinear, and therefore, to deliver correct 3D information, we must first accomplish stereo matching and calibrate the disparity map properly. Knowing the relationship between the two factors in a correct figure is indispensable. Accordingly, we build a standard scalar bar as in Fig. 2(a). There are 20 lines, drawn by a laser, on both sides of an aluminum plate of 1000 μm thickness. The distance from the ends to the center is divided into 50 μm increments to show differences of depth. Fig. 2(b) is an overlap image of a stereo image taken using an *in vivo* multiple-baseline system on the left side of the standard scalar bar. When we assume that the line showing the least depth between the object and the camera is at 0 μm, we can find the difference in disparity to the line at 1000 μm. In other words, we can confirm that the disparity increases as the line comes closer to the camera line.

In the right and left images photographed by this apparatus, the disparity between corresponding points is calculated using pixel counting. With this result, the formula relating disparity and depth can found by curve fitting. We can expect the 3D information to be closer to actual data by combining this formula and existing results.

3. Results

Fig. 3(a) is a picture of living skin obtained using an *in vivo* multiple-baseline stereo imaging system. The stereo image was obtained by positioning the main body close to the living skin. The cameras were set to produce TIFF images with a resolution of 1032 × 779 pixels at 24-bit color depth and the best available image quality. The field of view (FOV) adoptable to obtain the image was 5 × 3.8 mm at minimum and 80 × 60.4 mm at maximum.

Fig. 3(b)–(c) is *in vivo* stereo images of skin wrinkles obtained by this system. Fig. 3(b) is the overlap image of the right and left images obtained when baseline was 0 (OL = OR) by aligning the optical axes of each camera. It shows a clear skin feature that seems to come from one camera, although two images are actually overlapped. This shows that this system can obtain *in vivo* images without being affected by minute movements caused by breathing or other movements by obtaining right and left images simultaneously. In Fig. 3(c), which was obtained by changing the baseline to 76.2 μm, we see a blurred image caused by the disparity brought about between the right and left images. In other words, this shows that this system can obtain stereo images at the desired baseline. However, it is difficult to judge the depth of skin wrinkles with