

# Measuring Infant's Length with an Image

Maolong Tang\*, Ming-Ting Sun\*, Leonardo Seda\*\*, James Swanson\*\*\* and Zhengyou Zhang\*\*\*\*

\*University of Washington, Seattle, WA, USA, E-mail: {mltang, mts}@uw.edu

\*\*Universidade de São Paulo, São Paulo, State of São Paulo, Brazil, E-mail: leo.seda@hotmail.com

\*\*\*University of California, Irvine, CA 92697 USA, E-mail: jmswanso@uci.edu

\*\*\*\*Robotics X, Tencent, China, E-mail: zhengyou@tencent.com

**Abstract**— It is important to measure an infant's length regularly to estimate the growth velocity to make sure that the infant is growing normally. Traditionally, measuring an infant's length is performed with an infantometer. However, the infant struggles and cries in the measuring process, and it often needs three persons to position the infant's head, legs, and the boards of the infantometer during the process. Thus, it is not practical for a parent to perform this measurement at home regularly. In this paper, we propose a new approach which allows the measurement of an infant's length using a cellphone picture without the need to position the infant. Our algorithm automatically calculates the 3D positions of the body parts and the total length of the infant with the help of round stickers. The round stickers can be put on the infant's body easily in a few seconds, before the picture is taken. This new technology would make frequent measurements of the infant's length and the tracking of the growth velocity possible.

## I. INTRODUCTION

It is important to measure an infant's length and growth velocity regularly to make sure that the infant is growing normally. Traditionally, measuring an infant's length requires performing a manual measurement using an infantometer (which consists of a flat surface with a headboard and an adjustable footboard). The infant is placed on the flat surface, and then the head and the legs are straightened so that both touch the boards that are parallel, providing a measure of the infant's length. However, the infant struggles and cries in the process, and it often needs three persons to position the head and the legs of the infant, and to adjust the footboard of the infantometer. During the process, it could also accidentally hurt the infant. Fig. 1 shows an example of the process of measuring an infant's length using an Infantometer [1].



Fig. 1 Measuring the infant's length using an infantometer.

Given the complexity of the infant length measurement procedure and the staff requirement, a parent cannot be expected to perform this process alone regularly at home. It is very desirable to have a simple process so that the infant's length could be measured easily by a single person without special training or special equipment and the need to hold and straighten out the infant. One possibility is to use a photograph taken by a typical cellphone camera with the infant free-moving, and to use computer vision techniques to automatically calculate the infant's length by an algorithm.

However, this is a very difficult problem. A free-moving infant is a dynamic deformable object. The body parts of the infant (e.g., the head and legs) can bend in the 3D space. Also, the lengths of the body parts will appear differently in the 2D image depending on their distances and viewing directions relative to the camera. In order to accurately measure the length of the infant, we need the 3D positions of the joints of the infant's body parts relative to the camera, and physical reference lengths at these 3D positions since the same physical length could appear very differently in 2D images when at different 3D positions. We also need to be able to detect these body part joints accurately and robustly under various lighting conditions. Unfortunately, many body parts important for determining the length of the infant, such as the top of the head (which may be covered by hair), the neck, the shoulder, the joints between the body and the legs, the knee, and the heel, all lack clear and unique feature points. Another severe problem is that infants' movements may cause severe motion blurs of the body parts and joints in the image.

An intuitive approach to measure the length of an object in the 3D space from images is to construct the 3D structure of the object. With the 3D structure, the 3D coordinates of the feature points and the total length of the object can be calculated. The most popular technique to estimation the 3D structure from 2D images is Structure from Motion (SfM) [2]. SfM requires a rigid textured object, and the availability of multiple photographs from different angles. However, an infant is not a rigid object, and it is difficult to extract enough features from the relatively textureless skin. There are two main methods to handle non-rigid objects: one is the template-based method and the other one is the non-rigid structure from motion (NRSfM) [3] method. The template-based method usually requires not only the texture but also a known template, which makes it not applicable for our goal. The NRSfM method does not require a known model, however, this method requires accurate matched feature pairs [3]. Reliable feature matching such as SIFT [4] does not work well on skin with no apparent

texture. Dense matching by the state of the art optical flow algorithms [5] and [6] give very noisy feature pairs, so they are not suitable for accurate infant length estimation. Structure from shading is a method that does not rely on feature matchings. However, due to unknown light sources and the complexity of cloth material and skin, even the most recent structure from shading algorithm [7] cannot produce satisfactory 3D results for infant photographs.

In this paper, we aim to measure an infant’s length accurately and automatically from an image that is captured by a regular cellphone camera. In our proposed approach, we use round stickers of a known-size to mark the feature points (joints of the body parts) and to provide the reference lengths at the 3D positions of those feature points. A round sticker in the 3D space will be projected into an ellipse with different sizes and shapes depending on its 3D location and orientation relative to the camera. We can thus estimate the 3D positions of the centers of the stickers from the ellipses in the 2D image, and calculate the distances between the centers of the stickers on the infant. We develop a minimal spanning tree method to automatically determine the order of the ellipses for calculating the total length of the infant. We also developed a blurred-ellipse detection algorithm which can automatically reject the images containing ellipses with motion blur to obtain accurate results. The stickers can be put on the infant easily in a few seconds, and the calculation of the length is fully automatic.

We tested our algorithm on a model (a baby doll) as well as on human infants. The results show that our proposed approach can provide very accurate results and is easy to apply in practical situations.

The contributions of this paper include: 1. We propose a new approach which enables automatic calculation of an infant’s length from a picture taken from a regular cellphone camera. 2. We propose to use a minimal spanning tree based algorithm to automatically calculate the infant’s length. 3. We propose a blurred-ellipse detection method to reject the pictures with motion-blurred ellipses caused by the infant’s motion while taking the pictures. 4. We perform experiments to show the effectiveness and the accuracy of the proposed methods.

The organization of the rest of this paper is as follows. In Section II, we describe our proposed approach. In Section III, we discuss our proposed method to detect blurred ellipses due to motion, so that we could automatically reject the images containing blurred ellipses in our calculations in order to obtain accurate results. In Section IV, we present our measurement results. In Section V, we give conclusions.

## II. PROPOSED APPROACH

For a deformable object such as an infant, determining its 3D length from a single 2D picture is an ill posed problem due to the reasons mentioned above. Our idea is that feature points are needed, but they do not need to be “natural” feature points on the infant. We proposed to use known-size round stickers placed on the joints on the infant’s body to serve as easily identifiable feature points and provide the needed reference physical lengths at the 3D positions of those feature points. With different distances and viewing angles from the camera,

the round stickers will appear as ellipses in different sizes and shapes. By analyzing the ellipses’ geometric parameters, we can recover the sticker centers’ 3D positions relative to the camera with only one image. Stickers can easily be put on the infant’s joints (e.g., heel, knee, hip, shoulder, and head) in just a few seconds, and with the 3D positions of the ellipse centers, the distance between stickers could be calculated. The infant’s length could then be calculated by adding up the length of each body part. Since the hair may cover the infant’s head, we put a sticker on a hard surface (such as a cardboard or a book) to be pressed against the head to provide the length measurement between the top of the head and the shoulder.

With this approach, only off-the-shelf round stickers are needed in the measurement. There is no need to try to detect feature points in the textureless areas, or to provide reference physical lengths at the 3D positions of those feature points. Essentially, the stickers solve the problems of defining feature points, detecting the 3D locations of the feature points, and providing the physical reference lengths at the 3D locations all at once.

It was shown in [8] that the 3D position of the center of a round pattern relative to a camera could be determined up to two solutions based on the actual size of the round pattern and the parameters of the corresponding ellipse in the image. Those two solutions will have different surface orientations, but very close positions. Specially designed pattern is needed to help to resolve the ambiguity [9], otherwise, multiple images are needed to help to find the true position. For our situation, the round sticker’s size is very small compared to its distance to the camera center. Thus, there is no need to differentiate those two solutions: for a sticker of diameter 1.905 cm, camera to sticker distance larger than 20 cm, the difference of those two solutions is less than 0.5 mm. The 3D distance accuracy will not be affected by the sticker’s orientation, it is mainly affected by the accuracy of the detected ellipse parameters.

To measure the infant’s length automatically, we need to determine the connecting order of the ellipses automatically. To do this, we use a Minimal Spanning Tree (MST) based algorithm to obtain the connectivity of the stickers.

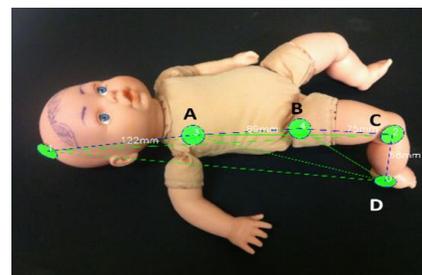


Fig. 2. Placing round stickers on the infant and using a minimal spanning tree based algorithm for calculating the infant’s length from an image.

We formulate the connections of stickers as a graph, in which each sticker represents one node. Let  $N_i$  be the  $i$ th sticker. Since any two stickers are connected by exactly one

path in our configuration, we can extract an MST to obtain the connectivity of the stickers. The optimal MST encourages that the path between any two connected stickers should be covered by the infant’s body as much as possible. For example, in Fig. 2, we encourage the connection between Node C and D, while discouraging the connection of A-D and B-D. Therefore, the path between two stickers will be weighted by the portion of the path located on the body as well as the Euclidian distance between the two stickers in 3D. More formally, the weight between any two nodes  $N_i$  and  $N_j$  is defined as:

$$E(i, j) = \|N_i^{3d} - N_j^{3d}\| \cdot (1 + \alpha p) \quad (1)$$

$$p = \frac{|Pixels\ outside\ the\ body|}{|Pixels\ along\ line(N_i, N_j)|} \quad (2)$$

where  $N_i^{3d}$  stands for the 3D position of node  $N_i$  and  $\| * \|$  is the Euclidian distance between the two points.  $| * |$  stands for the cardinal number, and  $\alpha$  is a constant. To count the number of the pixels outside the body, we first segment out the infant’s body from the background based on the skin color. The total length of the MST is the infant’s body length. Fig. 2 shows an example where green lines are all possible paths, and the blue lines are the selected MST path. Besides this MST-based method, an alternative method is to take multiple pictures with each picture containing two stickers only, and the total length is the sum of the individual length in each picture.

### III. BLURRED STICKERS DETECTION

Sometimes the ellipses in the picture appear blurry due to the infants’ movements. The blurred ellipses will have an impact on the accuracy of the final results. To improve accuracy, we develop a blur detection method to automatically reject the images with blurred ellipses. Since we can take multiple images, we have enough clear ellipses for the measurement. To increase the possibility of getting good images while the infant is moving, we can use the burst mode to take pictures.

Although the topic of image blur analysis has attracted much attention, most previous works focus on detecting blurred images. In [10], Su et al. constructed a new blur metric: singular value feature, and use it to detect the blurred regions of an image. In [11], Liu et al. designed four local blur features for blur confidence and type classification. In [12], Shi et al. studied a few blur feature representations in image gradient, Fourier domain, and data-driven local filters. In our case, even if there are some blurred regions in an image, as long as a sticker in the image is clear, it can still be used to calculate the 3D distance.

The motion blur could be modeled as a convolution with a motion blur kernel. The stickers are put on various backgrounds which have very different pixel values. Due to the convolution and the different pixel values across the ellipse boundary, the regions immediately inside and outside the ellipse boundary usually have relatively large gradients. Since the stickers could be put on many different backgrounds, the gradients outside the ellipse boundary have many variations.

On the other hand, since the inside of the ellipse is just pure green color in our case, we can ensure that the gradients inside the ellipse are small and consistent except the region immediately inside the ellipse boundary. Thus, it is easy to extract a ring with relatively large gradients between the ellipse boundary and the inside region of the ellipse. Depending on the width of the ring, we could determine the extent of the sticker blurriness.

After we extract the sticker masks by setting a color threshold, we calculate the magnitudes of the gradients of the grey-scale image for the pixels inside the sticker masks. The Sobel operator is used in the calculation of gradients. After that, we process the gradient images by binarization with an empirical threshold, followed by the Closing operation using a 9x9 kernel, which is an all-ones matrix, to fill the holes. An example for a blurred sticker is shown in Figure 3.

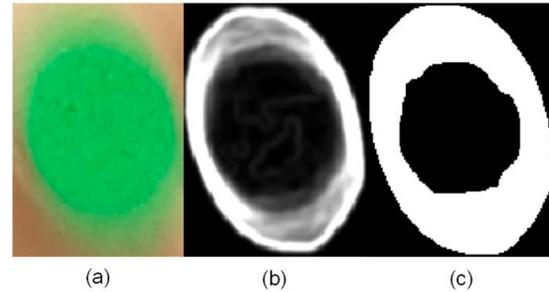


Fig. 3. Blurred sticker sample. (a) is the original blurred sticker image, (b) is the gradient image, and (c) is the mask of the gradient image after binarization and morphology processing.

To make the process simple, we approximate the gradient mask (e.g., in Figure 3(c)) by an elliptical ring, which is the blue part shown in Figure 4. We define the inner ellipse as the internal contour of the ring, the outer ellipse as the exterior contour of the ring. If the elliptical ring is wider, the corresponding sticker is more likely to be blurred.

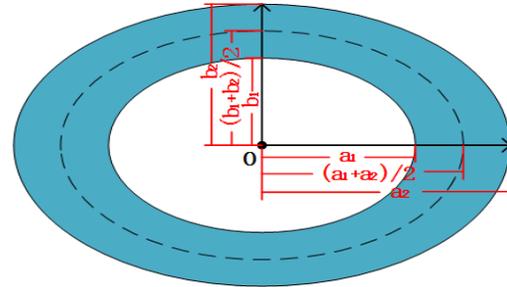


Fig. 4. The elliptical ring.

Suppose the long axis of the inner and outer ellipse are  $a_1$  and  $a_2$ , and the short axis of the inner and outer ellipse are  $b_1$  and  $b_2$ , respectively, where  $a_1 \geq b_1 > 0$ ,  $a_2 \geq b_2 > 0$ ,  $a_2 > a_1$  and  $b_2 > b_1$ . The middle dashed ellipse can be regarded as the skeleton of the elliptical ring, which can be obtained by skeletonization [13] of the elliptical ring, and the

long and the short axis of it are  $(a_1 + a_2)/2$  and  $(b_1 + b_2)/2$ . Thus the area  $A$  of the elliptical ring is

$$A = \pi(a_2b_2 - a_1b_1) . \quad (3)$$

The perimeter  $P$  of the dashed ellipse is

$$P = \pi(b_1 + b_2) + 2(a_1 + a_2 - b_1 - b_2) , \quad (4)$$

and the Area/Perimeter ( $A/P$ ) ratio is

$$A/P = \frac{\pi(a_2b_2 - a_1b_1)}{\pi(b_1 + b_2) + 2(a_1 + a_2 - b_1 - b_2)} . \quad (5)$$

To simplify the equations, we denote  $a_1 = n b_1$ ,  $a_2 = n b_2$ ,  $a_2 = m a_1$  and  $b_2 = m b_1$ , where  $n \geq 1$  and  $m \geq 1$ , so the  $A/P$  ratio in Equation (5) can be simplified into

$$A/P = \frac{(m-1)n\pi b_1}{\pi + 2(n-1)} \quad (6)$$

In Equation (6), for a given elliptical ring,  $n$  and  $b_1$  are fixed. If  $m$  is larger, the elliptical ring is wider.  $A/P$  estimates how thick the ellipse ring is. For accurate ellipse estimation, smaller  $A/P$  is preferred.

In our experiments, we tested our ellipse blur detection method on 30 infant images from the dataset we collected in our field trials. To evaluate the accuracy, we first manually estimated the stickers and classified the stickers as clear stickers or blurred stickers. First, we classified the images. If there were only clear stickers in an image, the image was classified as a clear image. Otherwise, the image was classified as a blurred one. Among the 30 images we tested, 15 images were clear and 15 images were blurred. Next, we classified the stickers. There were 86 stickers in these 30 infant images. If the Blur Ratio > threshold, the sticker was regarded as blurred; otherwise, the sticker was regarded as clear. We set the threshold as 20. The accuracy for the sticker classification was 95.3%, and only 4 stickers were incorrectly estimated, which may be due to the illumination and shadow effects.

Although four stickers were incorrectly classified, the 30 infant images tested were all correctly classified, so the accuracy of image classification was 100%. This is because the image is classified as a clear image only if all the stickers are classified as clear.

#### IV. EXPERIMENTAL RESULTS

##### A. Calibration of the Proposed Method as a Virtual Ruler

Before comparing the measurements of length using the proposed method and using an infantometer, we established the precision of the sticker method. For this purpose, we put six stickers on an Infantometer. Three stickers were put on the headboard to determine the headboard plane, and the other three at 20cm, 30cm, and 40cm positions on the flat base, respectively. Here the sticker's correct orientation out of two possible solutions could be selected automatically using the fact that the headboard stickers and the ruler stickers are

coplanar. By selecting the correct solution, the maximum 0.5 mm error mentioned in Section II can be circumvented.

We measured the distance from each of the three stickers on the headboard plane to each of the three stickers on the base of the infantometer to estimate the precision of our proposed method. As shown in Table 1, our proposed method is very accurate. The maximum measurement error is 0.11 cm, and the average is 0.058 cm (below 1 mm). This demonstrates that with clear images and precise ellipse detection, our proposed approach could achieve very high accuracy.

TABLE 1  
LENGTH MEASUREMENT (IN CM) USING STICKERS COMPARED TO INFANTOMETER, THREE IMAGES WERE TAKEN FROM DIFFERENT ANGLES.

Sticker Location	Image No	Dist2Headboard	Err	Err%
20 cm	1	19.89	0.1	0.6%
	2	19.98	0.02	0.1%
	3	19.97	0.02	0.1%
30 cm	1	29.92	0.08	0.3%
	2	29.96	0.04	0.1%
	3	30.07	0.07	0.2%
40 cm	1	39.94	0.06	0.15%
	2	40.03	0.03	0.08%
	3	40.11	0.1	0.3%

Next, to evaluate the accuracy in a controllable environment, we put stickers on a baby doll's joints. We used a book touching the doll's head. The length between the head to the shoulder is estimated by calculating the shoulder sticker's distance to the book plane. In this baby doll experiment, only one image is used to estimate the total length, since all stickers were captured in one image. One person could take the pictures easily.

The final infant length is:

$$L_{total} = L_{head2shoulder} + L_{shoulder2hip} + L_{hip2knee} + L_{knee2heel} + r_{sticker} \quad (7)$$

As shown in Table 2, the final result has only about 0.97% error compared to the ground truth measured by an infantometer.

TABLE 2  
BABY DOLL MEASUREMENT WITH 3/4" GREEN STICKERS. (RESULT IN MM), GROUND TRUTH (GT) OBTAINED BY REAL RULER.

NO.	Head2Br east	Breaset2 Hip	Hip2Knee	Knee2Heel	Total
1	111.55	95.2	73.83	65.83	346.41
2	111.11	94.40	71.91	65.69	343.11
3	113.67	95.09	73.03	65.14	346.93
4	114.47	96.71	73.41	63.91	348.50
5	112.03	97.18	72.19	64.43	345.83
6	111.09	96.44	72.21	67.28	347.02
7	112.15	94.46	74.62	66.83	348.06
8	112.91	95.44	73.29	65.32	346.96
AVG	114.24	95.50	73.00	65.49	348.24
STD	1.22	1.04	0.93	1.13	1.64
GT	114	96	73	67	350
diff%	1.43%	0.40%	0.084%	2.16%	0.97%

B. Field test

We also apply the proposed method to real infants in practical settings. Since it is more difficult to take a single clear picture that contains all stickers, we take pictures of knee to heel, hip to knee, shoulder to hip, and head to shoulder. Different types of pictures were uploaded to different folders and the computer automatically calculates the individual length and sums the individual lengths to get the final total length.

The size of the sticker in the image can also affect the accuracy of the measurement. Although theoretically we could recover circular stickers of any orientation and distance, the result depends on the accuracy of the ellipse parameters. If a circular sticker is too small in the image, then even a small deviation in the ellipse estimation will cause noticeable errors in the 3D position recovery. Thus, it is preferred to use a reasonably large sticker. However, if the sticker is too large, it is not convenient to use. We used a 3/4" (1.905 cm) sticker, which is a suitable size. Also, we made the stickers appear large in the pictures taken.

Another issue is the location of the stickers. For example, a sticker could be put on the outer side of a leg or on the inner side of the leg. The principle is to place the stickers close to the joints and observe if the distance between two neighbor sticker pairs is affected when the baby moves.

We compared the infant length measurements using both the proposed method and the infantometer method of the same day (see Table 3). We averaged the results if there were more than one clear photo available. The errors are close to 3%.

TABLE 3  
FIELD TEST COMPARED WITH INFANTOMETER. (IN CM). STICKER DIAMETER IS 1.905 CM.

Sub	Head to Shoulder	Shoulder to Hip	Hip to Knee	Knee to Heel	Total	GT	Err
1	16.77	18.30	11.88	13.98	61.89	60.0	3.15%
2	16.74	20.44	15.41	13.25	66.80	66.5	0.5%
3	19.7	31.46 *		14.27	66.38	65.0	2.1%
4	18.51	16.03	13.58	13.70	61.90	64.0	3.3%
5	15.00	19.23	11.92	13.43	59.65	64.3	7.2%
6	17.70	20.58	13.96	14.07	67.27	67.45	0.27%

\* Combined result.

V. CONCLUSIONS

In this paper, we propose a new technique to measure an infant's length from one 2D photograph. The method is easy to apply, and requires only circular stickers and a regular cellphone camera. Thus, a parent could easily apply this method regularly to keep track of the infant's growth. Compared to the traditional infantometer method, this new method can be applied by one person (instead of three persons). It also does not need special equipment (i.e., an infantometer). Our experimental results show that it has a good accuracy of about 3% in field trials. Since the major error is due to the head measurement and the sticker locations, it is possible to further improve the accuracy and repeatability with clearer definitions of head measurement and the joint locations.

ACKNOWLEDGMENT

This project was partially supported by Gates Foundation, grant number # OPP1142172.

REFERENCES

- [1] M. de Onis, A. W. Onyango, J. Van den Broeck, C. W. Chumlea and R. Martorell, "Measurement and standardization protocols for anthropometry used in the construction of a new international growth reference," *Food and nutrition bulletin*, vol. 25, no. 1 suppl1, pp. S27-S36, 2004.
- [2] R. Hartley and A. Zisserman, *Multiple view geometry in computer vision*, Cambridge university press, 2003.
- [3] M. Salzmann and P. Fua, "Deformable surface 3d reconstruction from monocular images," *Synthesis Lectures on Computer Vision*, vol. 2, no. 1, pp. 1--113, 2010.
- [4] D. Lowe, "Distinctive image features from scale-invariant keypoints," *International journal of computer vision*, vol. 60, no. 2, pp. 91--110, 2004.
- [5] D. Sun, S. Roth and M. J. Black, "Secrets of optical flow estimation and their principles," in *Computer Vision and Pattern Recognition (CVPR)*, 2010.
- [6] Q. Chen and V. Koltun, "Full Flow: Optical Flow Estimation By Global Optimization Over Regular Grids," in *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, 2016.
- [7] J. T. Barron and J. Malik, "Shape, illumination, and reflectance from shading," *IEEE transactions on pattern analysis and machine intelligence*, vol. 37, no. 8, pp. 1670--1687, 2015.
- [8] Q. Chen, H. Wu and T. Wada, "Camera calibration with two arbitrary coplanar circles," in *European Conference on Computer Vision*, 2004.
- [9] A. Pagani, J. Koehler and D. Stricker, "Circular markers for camera pose estimation," in *WIAMIS 2011: 12th International Workshop on Image Analysis for Multimedia Interactive Services*, Delft, The Netherlands, 2011.
- [10] B. Su, S. Lu and L. C. Tan, "Blurred image region detection and classification," in *Proceedings of the 19th ACM international conference on Multimedia*, 2011.
- [11] R. Liu, Z. Li and J. Jia, "Image partial blur detection and classification," in *IEEE Conference on Computer Vision and Pattern Recognition*, 2008.
- [12] J. Shi, L. Xu and J. Jia, "Discriminative blur detection features," in *IEEE Conference on Computer Vision and Pattern Recognition*, 2014.
- [13] T. Y. Zhang and C. Y. Suen, "A fast parallel algorithm for thinning digital patterns," *Communications of the ACM*, 1984.