Computer Vision and Image Understanding 000 (2016) 1–12

Contents lists available at ScienceDirect

Computer Vision and Image Understanding

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

Toothbrush motion analysis to help children learn proper tooth brushing

Marco Marcon*, Augusto Sarti, Stefano Tubaro

Dipartimento di Elettronica, Informazione e Bioingegneria, Politecnico di Milano, p.zza Leonardo da Vinci, 32 20133 Milano, Italy

A R T I C L E I N F O

Article history:
Received 24 April 2015
Revised 16 March 2016
Accepted 18 March 2016
Available online xxx

Keywords:
Target tracking
Hidden Markov models
Color segmentation

A B S T R A C T

Toothbrush training is a complex and not fun task for the child nor for the parents or for the dental stuff. Parents and hygienists often report that they are frustrated by poor responses to the training and in most of cases children go home and resume wrong brushing habits, if any. In this paper we present a novel approach where the tooth brushing procedure can become a fun and enjoyable task for kids using a cheap toothbrush accessory and a tablet or a smartphone. The main idea is to apply a simple and cheap 3D colored target at the end of the toothbrush and to track and analyze its motion, imparted by the child. In particular, from the tablet camera it is possible to track both the toothbrush target and the child's facial parts in order to estimate the brushed dental side. The proposed approach has been tested on seven kids showing good results both in propensity and accuracy after a 20 days period.

1. Introduction

Thanks to the widespread of handheld powerful devices, there is a plenty of opportunities where they can permeate our daily life with novel and helpful applications. In particular, thanks to the significant computational resources and the large number of sensory input/output devices they fulfill the mobile cyber-physical systems paradigm of intelligent embedded systems (Alippi, 2014; Dziri et al., 2014; Yoon and Ahn, 2006). In this article, we describe a novel application where mobile devices can fruitfully perform as an oral hygiene supervisor for kids. The usage of integrated cameras, high computational power and handless user interaction make these devices the perfect tooth-brushing assistant in order to make the procedure interactive, stimulating and fun. Most of kids find tooth brushing an annoying procedure and even if their parents/caregivers spend a lot of time trying to teach them the correct oral hygiene, they can easily acquire wrong habits. A correct tooth brushing training requires a constant supervision and the ability to create stimulating tricks to make children greet with enthusiasm. In order to simplify training and supervision tasks, many companies such as Oral-B (2014) and Kolibree (2014) are introducing new electric toothbrushes; these devices, equipped with Bluetooth connectivity, gyroscopes and accelerometers, communicate with smartphones or tablets capable of registering all the toothbrush motions in order to keep track of the brushed dental parts and the time dedicated to each of them. Electric toothbrushes equipped with advanced sensors cost significantly more than a manual toothbrush, they require to be charged or to replace batteries and, even if they could be more fun and entertaining, it is easy to accidentally break them: dropping an electric toothbrush can be fatal. Furthermore, kids could try to cheat the system just mimicking the tooth brushing gestures keeping the toothbrush head out of their mouths since no visual analysis is performed.

In order to keep the whole system for tooth brushing analysis as cheap and simple as possible, we decided to track the toothbrush and the face of its user, by the smartphone frontal camera. The smartphone display plays the role of a "*virtual mirror*" where the image of the person using the toothbrush is replaced by an avatar that, thanks to the facial features tracking, is able to completely mimic his gestures and expressions and points out wrong movements. The toothbrush itself can be replaced by a virtual tool and an interactive videogame can be realized where experience points and level progression are steered by a proper brushing procedure and timing. In our approach, we simply evaluated the effect of an avatar that mimics the children movements and correct their wrong gestures or timing on each side of the dental arches. We also evaluated in Section 5.1 the results of removing this tool after 10 days of usage to analyze the learning effect independently from the app usage.

The most relevant aspects, related to Computer Vision and Pattern Analysis, regard the tracking of the toothbrush, the

* Corresponding author. Tel.: +390223999639; fax: +390223999611.
E-mail addresses: marco.marcon@polimi.it (M. Marcon), augusto.sarti@polimi.it (A. Sarti), stefano.tubaro@polimi.it (S. Tubaro).

http://dx.doi.org/10.1016/j.cviu.2016.03.009
1077-3142/

estimation of its spatial position and orientation with respect to the user mouth and the analysis of its motion.

The toothbrush spatial position and orientation is obtained applying a colored target at the end of the toothbrush handle: this is the only portion of the toothbrush that will always be visible during tooth brushing since the head could be inside the mouth or could be covered by toothpaste and foam while the user's hand, grasping the handle, will cover it. The colored target is a ball in order to avoid dangerous sharp corners and is painted with vivid colors that can be easily segmented from the background using a suitable color space. In literature we found some other approaches based on target tracking for toothbrushing analysis, in particular Flagg et al. (2011) adopted color segmentation for toothbrush localization together with face parts localization but in this preliminary work a simple planar localization is considered without any 3D rotation or translation tracking or analysis in Chang et al. (2008) a shining marker equipped with LEDs is adopted on the tail of the toothbrush but no movements analysis is performed (a frame by frame analysis is performed), furthermore the mouth of the user is not tracked so it is impossible to know if the child is brushing his teeth or just moving the toothbrush in the air. Most of other approaches Lee et al. (2007); 2012); 2011) are based on accelerometers, magnetometers and/or gyroscopes which represent a different approach to the problem requiring instrumented toothbrushes with batteries and transmission interfaces. Furthermore bulky and heavy devices do not appear particularly appealing to young people for improving their inclination towards tooth brushing (Strickland, 2013). In our approach the idea is to reuse as much as possible of the computational power and acquisition capabilities of common tablets and smartphones and, at the same time, with the same hardware, we can also monitor facial parts position getting a more accurate estimation of the brushed dental side. Results with other approaches are compared in Section 5. In our approach the target tracking is then performed using a modified version of the MOSSE (Bolme et al., 2010) algorithm based on correlation filters: the aim is to account for the blurring due to the rapid motion of the target. In fact the toothbrush during the brushing procedure is usually subject to fast and periodic translations and rotations. The analysis of the toothbrush motion is performed using a Hidden Markov model (Bashir et al., 2005; Rabiner, 1989), this will allow us to estimate motion paths imparted by the user.

In order to evaluate the correctness of the brushing procedure the user face and its parts are also tracked; in particular we used the open source Active Shape Model code (Xing, 2010) in order to get an efficient estimation of eyes, nose and mouth position. Using this information we will estimate which surface of the dental arches is being brushed and in which direction.

The article is articulated as follows. In Section 2 we describe the target and how we can localize it inside an image through color segmentation. Section 3 describes the tracking procedure adopted for blurred images. Section 4 shows the procedure to recognize the dental arch brushed surface and evaluate the correctness of the brushing motion. The article concludes with Sections 5 and 6 where results on a set of children are presented and commented.

2. Toothbrush localization and characterization

The first step for the tooth brushing analysis consists of a robust toothbrush tracker, in order to get this an easy-to-track target was adopted: we used a ball with vivid and highly saturated colors in order to easily segment and localize the target from the background even in unconstrained and poorly illuminated environments. Many techniques were proposed for robust color segmentation in noise environments and Fuzzy C-means (FCM) has been widely adopted (Kim et al., 2004) together with its variants (Xess and Agnes, 2013) like SFCM (Spatial FCM) (Chuang et al., 2006) or THFCM (Thresholding FCM) (Jassim, 2012). FCM approach introduced the fuzzy concept so that the same object can belong to more than a single class at the same time. For every object, the belonging degree to a class is related to the association strength between the object and the class. The largest drawbacks related to the FCM approach are that the number of clusters must be set before clustering, the resulting quality is strictly connected to initial clusters seeds and the overall approach is not spatial dependent. In our application, we can *a priori* define the target colors; the color drift in the acquired frames is due to different illuminants, shadows or different acquisition settings (white profile). The *a priori* knowledge of the approximate searched colors allow us to simplify the initial target localization and its further segmentation in the tracking phase. In particular we followed a supervised Region Splitting and Merging (RSM) approach starting from the unsupervised approach proposed in Tan et al. (2013). Accordingly to Agoston (2005), we adopted, as a color space the Hue, Saturation, Value (HSV); further details on this color space conversion can be found in Gonzalez and Woods (2008).

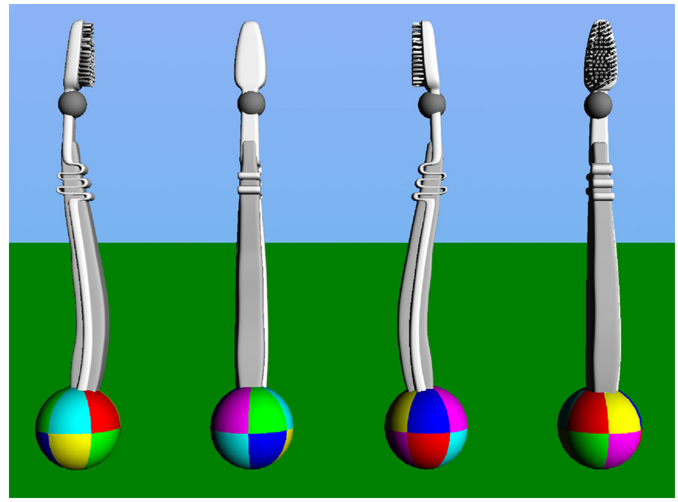

Fig. 1. The target colors seen from different angles.

2.1. Target description

In order to be able to estimate the toothbrush spatial orientation the target is painted with a specific color sequence. In particular, the target is composed of two hemispheres whose joining circle (the equator) is orthogonal to the main toothbrush axis. Every hemisphere is divided into six slices based on six colors chosen in order to be maximally distant in the Hue dimension: Red, Green, Blue, Cyan, Magenta,Yellow. The two hemispheres, to avoid ambiguities, where painted (see Fig. 1) following two different color sequences:

- the top hemisphere: Red, Yellow, Blue, Magenta, Green, Cyan.
- the bottom hemisphere: Green, Magenta, Red, Cyan, Blue, Yellow.

The target colors present a high saturation and are painted on a dull surface in order to prevent highlights. The first screening to localize possible pixels belonging to the target is based on a Value threshold and a Saturation threshold.

2.2. Target localization

In real applications, where the environment is usually a bathroom and the user keeps the smartphone or the tablet in front of him, the light sources are quite intense and are placed above