Feedback Mechanisms of an iOS App to Record RGBD Data for AI-Based Diagnosis and Monitoring of Infant Head Deformation

Fabian RAPP *[1], Samuel ZEITVOGEL [3], Christian MEDER [1], Kai STOEVESANDT [2],
Christian WERNET [3], Astrid LAUBENHEIMER [3]
[1] inovex GmbH, Karlsruhe, Germany;
[2] VARILAG GmbH & Co. KG, Ettlingen, Germany;
[3] Intelligent Systems Research Group (ISRG), Karlsruhe University of Applied Sciences, Germany

https://doi.org/10.15221/22.39

Abstract

We present a prototype of a 3D infant head reconstruction app on a mobile phone for a cranial asymmetry measurement method. In this work, we focus on the user experience and user feedback of the interactive reconstruction pipeline.

The advent of 3D sensor-equipped mobile phones such as the iPhone TrueDepth camera enables widespread 3D reconstruction and measurement applications such as room measurements for interior design preview and object reconstruction. Reconstructing an infant's head with an RGBD sequence is challenging due to the uncooperative behaviour of the subject.

The RGBD data for the scan is recorded with the TrueDepth camera system on the front of iPhones (starting with the iPhone X), which is normally used for FaceID. The forward-facing position of the sensor poses a challenge to the app operator because the screen is hardly visible during recording.

The major contribution of this work is an app providing audible, haptic and visual feedback to support the user during the scanning process. Furthermore, the app is designed to be easy to use following an intuitive user experience design. Lastly, the developed app is tested by a team of physical therapists to evaluate the feedback concept and the user experience.

Keywords: 3D, TrueDepth camera system, user feedback, CVAI measurement, scanning app

1. Introduction

Starting in the early 90s, it is recommended that infants sleep on their backs to mitigate the Sudden Infant Death Syndrome (SIDS). While keeping their infants on their backs successfully reduces the risk of SIDS [1], studies hint at an increase in an infant's head shape deformation (plagiocephaly), favoured by a homogeneous sleeping posture [2]. Plagiocephaly refers to the medical condition of the skull deformation caused by persistent external directional forces [3]. Studies indicate that 37 to 46 percent [4, 5] of infants 7 to 12 weeks of age suffer from plagiocephaly. While plagiocephaly can be successfully treated, unaddressed plagiocephaly can lead to a multitude of risk factors for the infant [6-11]. Depending on the severity of the plagiocephaly there are different treatment possibilities ranging from a simple change of sleeping positions to specifically formed pillows to helmet therapy [12]. The most common technique to diagnose plagiocephaly is to use a calliper to measure the diagonal lengths of each side of the head and use these measurements to calculate the Cranial Vault Asymmetry Index (CVAI) which defines the severity of the plagiocephaly. However, measurements taken by different examiners might vary by as much as 2.2 millimetres [12].

In this work, we present novel user feedback mechanisms for a 3D infant head reconstruction app on a mobile phone for a cranial asymmetry measurement method. The major contribution of this work is an app providing audible, haptic and visual feedback to support the user during the scanning process.

This paper starts with a brief overview of the conditions of a scanning procedure in section 3.1. Afterwards, the user experience concept is explained in section 3.2 by presenting the feedback possibilities of an iPhone in section 3.3, continued by the logic behind the feedback in sections 3.4 to 3.6. In the end, this feedback is tested and its evaluation is presented in section 4.

* fabian.rapp99@gmail.com

2. Related Work

There are various works about 3D reconstruction using a mobile phone [13, 14]. Barbero-García *et. al.* [15] propose a technique for creating a 3D model of a head by using a cap coded with markers that serve as reference points to the capture device. This technique is also used in the app of this work. However, only a few focus on the feedback aspect of the scanning procedure which is the focus of this paper.

Tanskanen et. al. [16] propose a feedback technique utilizing visual feedback with information from inertial sensors of a mobile phone combined with visual pose estimation and 3D dense modelling for ensuring quality and completeness for a 3D reconstruction pipeline.

Kim et. al. [17] use a Microsoft Kinect scanning device for indoor objects and provide visual feedback to the user to ensure quality and completeness of the scan. The visual feedback consists of showing the current point cloud of the scanned object and displaying the best matching database model to the currently scanned object with still missing parts highlighted in colour.

Locher et. al. [18] also use visual feedback as a means to optimize the 3D reconstruction result. Their setup includes a mobile device as a frontend and a cloud component as the backend. On top of displaying the current point cloud, they show a map that illustrates good spots where to take further images.

3. Method

This chapter will present the user experience concept of the app by explaining the reason why feedback is even necessary. Afterwards, the feedback possibilities of an iPhone are explored and we will present what data is used to provide feedback, how it is used and how it is presented to the user.

3.1. Scanning conditions

Fig. 1. A doll wearing a cap with markers that was used for testing.

For the scanning procedure, the infant has to wear a special cap (see figure 1) with ArUco markers [19] on it. Each ArUco marker serves as a reference point to be able to construct a 3D model from all the recorded frames. Barbero-García *et. al.* [15] presented this technique.

The goal of this app is to collect RGBD data of the heads of infants, so video data, combined with depth information. This information is used to construct a 3D model of the infant's head and to train a model. With a trained model the goal is to be able to create a scan without needing the infant to wear a cap in a future version of the app. The model training and 3D model reconstruction is not done in the app, so it does not have to be in real-time. The 3D model can also be used to calculate the CVAI.

For a successful construction of the 3D model, some aspects of the scanning procedure must be considered:

The video frames recorded by the camera must be all connected, meaning that each frame should have an overlap with its neighbouring frame. This overlapping section provides points that are present in both frames. These reference points are the connection between the frames and allow the construction of a coherent 3D model. This is the reason why the infant has to wear a cap with ArUco markers because these markers work as the described reference points.

To also calculate the CVAI from such a 3D model, it is further important that the 3D model also includes the tragion (a part of the ear), the nasion (the part between the top of the nose and the forehead) and the nose tip. Zeitvogel *et. al.* [20] provide a detailed explanation of how the CVAI calculation using these parts of the head works.

So, during the scanning procedure, the user has to make sure that there is sufficient overlap between the frames and that the mentioned parts of the head are clearly visible in the scan.

3.2. User Experience Concept

The purpose of this app is to create scans of the heads of infants with the help of the TrueDepth camera system of the iPhone.

Fig. 2. Sketch of an iPhone in landscape mode.

The camera system (see black area in figure 2) is located at the front of the iPhone at the top of the screen, so on the same side as the display. This location creates a challenge for the person creating the scan. To scan the infant's head the camera has to point to the infant which means that the camera and the iPhone screen point away from the app operator. However, even without properly seeing the iPhone screen, the user still has to make sure that the head of the infant is in the camera's view, meaning it is not too close, not too far away and not just partially in the camera's view.

The next challenge for the user is to somehow know when the entire head has been recorded so that the recording can be safely stopped. After all, the goal of this version of the app is to collect high-quality data to train a model. When it is time to train the model, each scan should have all the necessary data (as explained in section 3.1). That is why there has to be feedback to the user regarding the completeness of the scan.

3.3. What feedback output mechanisms does the iPhone offer?

The first step is to find out all the features that an iPhone offers that could be used to provide visual, audible and haptic feedback to the user.

3.3.1. Screen

The most obvious solution for giving feedback on a phone is the screen, as it offers a wide variety of visual feedback possibilities. But as described in the previous section, the screen is not always visible to the person recording as the phone faces in the other direction. Therefore, the screen can very well be used for any feedback before starting a recording and after stopping a recording, but it is not ideal for feedback during a recording.

3.3.2. Speaker

Another option is to use audio feedback. Audio output has the advantage that it doesn't matter which way the phone is turned, the audio just has to be loud enough for the user to hear it. Audio allows us to use simple sounds like in a parking sensor or use actual recordings of words or even sentences. Words or sentences have the advantage of being very descriptive. One thing to keep in mind is that the user of this app is dealing with infants who might react very differently to audio than the user would. A loud beeping sound like in a parking sensor could frighten the infant which can't be the goal of this app. One solution could be to wear headphones while recording so that the infant does not hear the sounds.

3.3.3. Vibrations

A further non-visual feedback mechanism is haptic feedback, i.e. vibrations. Vibrations are normally used to, for example, indicate that a new message arrived, that a timer has finished counting or that somebody is calling. Vibrations can be customized regarding their intensity, their sharpness and their duration. Just like with audio feedback there is also the advantage that the phone has certain default vibrations for success and error that the user is already familiar with. This could make the usage of the app easier when a user can already associate certain vibrations with certain meanings.

3.3.4. Flashlight

A final option is the flashlight of the phone. It is conveniently placed on the back of the iPhone which is visible to the user at all times during the recording. It does not offer as much variety as audio or haptic feedback, but it can still be controlled regarding its brightness and how long it is turned on. However, like audio feedback, it might scare the infant and it has to be examined whether the brightness is bearable for the user or if it is distracting.

3.4. What data is used to create feedback?

The ArUco markers on the cap that the infant is wearing are used as a data source for both the positioning and the progress feedback.

Each marker delivers multiple pieces of information:

From each corner of a marker, we get that corner's x and y coordinates in the recorded picture. Furthermore, we get the distance of that corner from the camera with the help of the TrueDepth camera system. The pattern of an ArUco marker represents a unique ID that identifies this marker.

On top of that, the app calculates the overlap status of the recorded frames so that a coherent 3D model of the infant’s head can be created later.

3.5. How data is used to determine the corresponding feedback

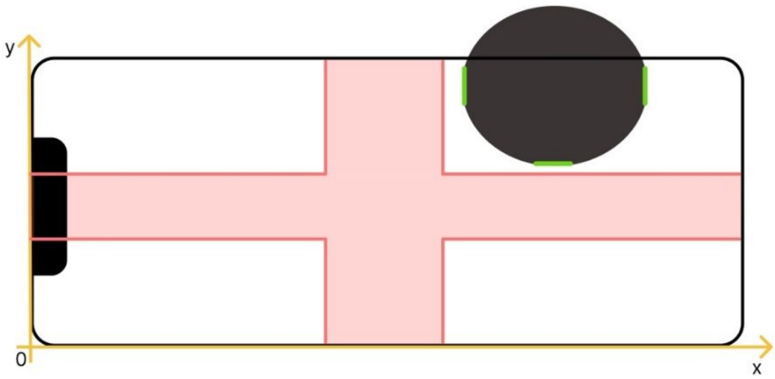

Fig. 3. Determine 2D position of the head within the recorded image

3.5.1. Positioning Feedback

The positioning feedback is divided into two parts. The 2D positioning and the distance from the head.

To get the 2D positioning, i.e. the positioning of the infant’s head in relation to the iPhone, first the head has to be detected in the recorded image.

To find the head in the image the app goes through all the x and y coordinates at least once per frame and determines the maximum and minimum values on both the x and y axis (see the green lines in figure 3). Furthermore, an area that counts as the centre is defined (see the red area in figure 3) as shown in the following equations.

The centre area on the x-axis is denoted by p and the height of the image by h.

$$\frac{h}{2} - \frac{0.05h}{2} \leq p \leq \frac{h}{2} + \frac{0.05h}{2} \qquad (1)$$

The centre area on the y-axis is denoted by q and the width of the image by w.

$$\frac{w}{2} - \frac{0.05w}{2} \leq q \leq \frac{w}{2} + \frac{0.05w}{2} \qquad (2)$$

With that, the position of the head is found and feedback can be determined accordingly. In the scenario, in figure 3 the feedback would be to move the device right-down to centre the head.

The second part of the positioning feedback is the distance of the iPhone from the head. The distance for each corner is determined with the TrueDepth camera system. When the device gets too close to a corner of a marker, invalid depth information is returned, so instead of a valid number the value "not-a-number" is returned.

3.5.1. Progress Feedback

In each recorded frame the app identifies the IDs of all the visible ArUco markers. Each marker has a unique ID. The number of scanned markers serves as a reference for the progress of the scan. To calculate the progress, the user first has to enter the number of ArUco markers on the cap in the settings view of the app. During the scanning procedure, the app counts the number of detected marker IDs and compares it to the total number of markers. Once all marker IDs have been seen in the scan, it is complete.

3.6. Combining data with a feedback mechanism

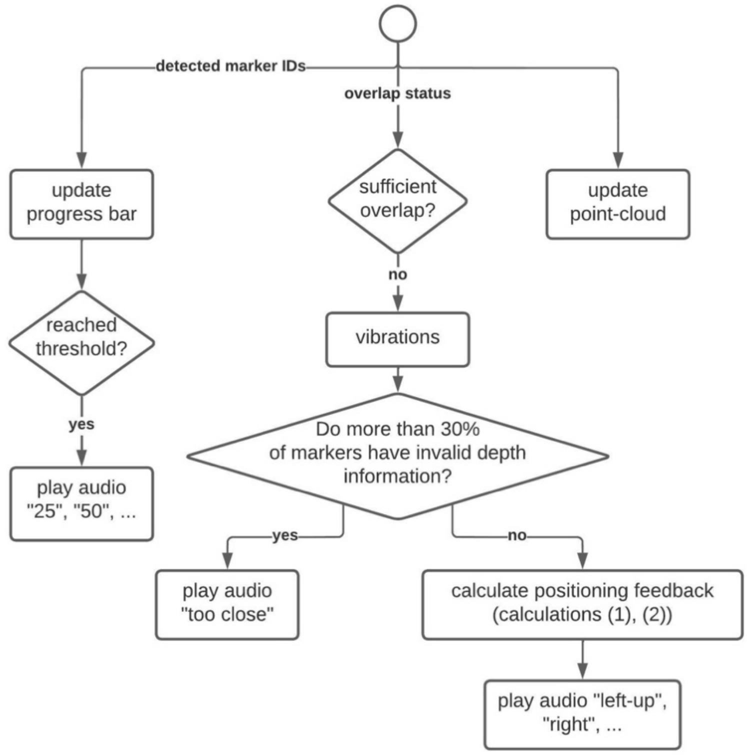

Fig. 4. Flow chart of the feedback concept. Rhombus shapes represent decisions and rounded rectangles represent actions that the app takes.

After calculating the feedback, this information has to be presented to the user in an intuitive way.
As evaluated earlier, using haptic and audible feedback appears to be the most efficient way to do this. Refer to figure 4 for a visual overview of the feedback concept.

3.6.1. Positioning feedback

- **Primary feedback:** The marker overlap status between the frames serves as the primary criteria for whether the positioning of the iPhone is satisfactory or needs to be improved. To signal to the user that something about the positioning needs to be corrected, the iPhone vibrates. This vibration is continuous and only stops once the overlap is sufficient again.
  However, the vibration is only started and stopped when the marker overlap is the same four frames in a row. This threshold is important as the app processes 30 frames per second and therefore, we need to reduce the number of times feedback is triggered.

- **Secondary feedback:** The positioning feedback regarding the distance or the 2D positioning is only triggered if the primary feedback has been triggered and provides further information to the user on how to correct the positioning.
    - Audio output "too close" if more than 30 percent of all markers have invalid depth (depth value = not-a-number) information. This threshold is derived heuristically from initial testing.
    - Audible positioning instructions ("left-up", "right", ...) if less than 30 percent of all markers have invalid depth information.

3.6.2. Progress feedback

- Determined by the number of detected marker IDs. The user can select in the settings view of the app between:
    - Audio output every 25 percent.
    - Audio output every 50 percent.
    - Audio output at 100 percent.
- The progress is also shown visually as a progress bar on the screen. The colors of the progress bar match the colors of a traffic light. This makes the understanding of the progress bar intuitive. This was implemented because of the scenario where the app operator does not hear the progress audio output because the infant cried for example. Then he/she can quickly check the screen.

4. Testing and Evaluation

To test and evaluate the features of the app and even more importantly the user experience of the feedback concept, it is used by a team of physical therapists in a four-week evaluation period.

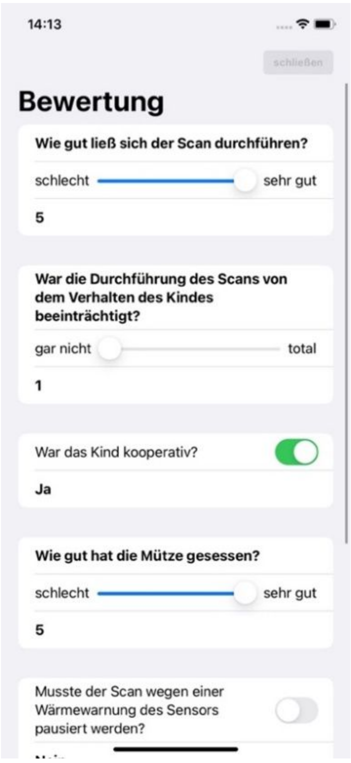

Fig. 5. Screenshot of the rating screen that is shown after every scan.

During that time, direct feedback about the scanning process is collected inside the app. This feedback is collected after every scan. For feedback collection, an evaluation form (see figure 5) consisting of sliders and toggle switches is developed to collect immediate feedback right after the scanning procedure. These questions deliver an initial impression of how the scan went and what the situation was like during the scanning procedure.

To get an even better understanding of the pros and cons of the app, an in-person interview is conducted with the physical therapists that used the app during the evaluation period. The results of both evaluations will be presented in this section.

The following questions were asked after every scan:

- How well did the scan go?
- Was the performance of the scan affected by the child's behavior?
- Was the child cooperative? (The child did not excessively turn his/her head, it did not try to pull off the cap and it did not put its hands in the camera's view.)
- How well did the cap fit?
- Did the scan have to be paused because of a sensor heat warning?
- Did visual confirmation need to be obtained in addition to audible and haptic feedback on the screen?
- The following feedback configurations were also saved in the rating file:
    - Play progress audio only at 100 percent.
    - Play progress audio every 50 percent or every 25 percent.
    - Play positioning audio only if the device is too close to the infant's head.

4.1. Results and evaluation of the scan ratings

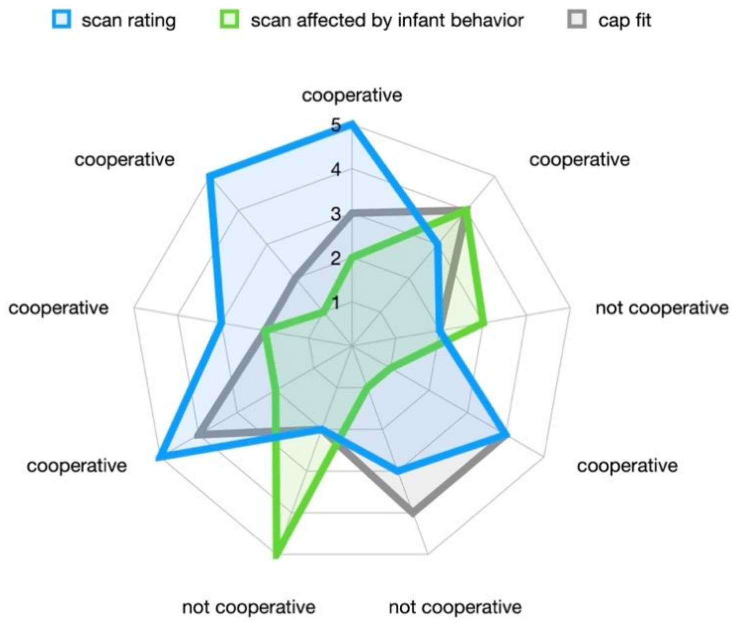

Fig. 6. Evaluation of the data collected inside the app after every scan

Nine scan ratings were collected in total.

Figure 6 shows a net diagram displaying the overall scan rating if the scan was affected by the behaviour of the infant and how well the cap with the markers fits the head. The diagram puts these metrics in relation to the cooperating behaviour of the infant.

This data suggests that the reason for bad scans (rating of 1 or 2) could stem from the infant's behaviour and the fit of the cap. However, good scans (rating of 4 or 5) do not seem to correlate with the cap fit and the infant's behaviour which suggests that the scan was good because the app fulfilled its requirements. Only one scan suggests that there might be an issue with the online 3D reconstruction pipeline.

Lastly, the average scan rating is 3.56. In total there were four very good (ratings of 4 or 5) scans and two bad (ratings of 1, 2) scans.

4.2. Results of the in-person interviews

After analysing the collected data, the interviews with the physical therapist can validate certain assumptions. In total, four physical therapists used the app, ranging from one to four times each.

4.2.1. Positioning feedback

All interviewees agreed that the part of the positioning feedback that provides information about the distance to the head via audio was really helpful and always allowed for an intuitive repositioning of the phone.

The opinion about the vibrations that indicate that something about the positioning is off, differed. For one interviewee it sometimes felt like the iPhone was constantly vibrating, but when looking at the screen and inspecting the displayed point cloud of the head, the positioning appeared to be fine. Another interviewee, however, liked them.

The positioning instructions on how to move the iPhone were quite similar among all interviewees. They found them a little unintuitive and sometimes hard to interpret. They didn't say that they appeared wrong, just that they were hard to put into action.

4.2.2. Progress feedback

One interviewee reported that the progress feedback was motivating for the parents holding their children.

Some interviewees wished that the audible feedback part offered smaller increments although the app does in fact offer these increments. Because the physical therapists didn't have much time for the scanning procedure, it was not possible for them to explore all features of the app.

4.2.3. Cap

The cap that was used is still a prototype and therefore the ArUco markers were only glued onto the cap. Still, all interviewees agreed that the cap was the biggest problem and the cause of a bad scan. First, the cap didn't fit equally good for all infants. For some infants, the closure of the cap was a little short. Second, some interviewees complained that the cap was too small for some children.

Third, other interviewees mentioned that the stickers did not hold well and would often fall off which caused an interviewee not to reach the scan completion in some cases because of the number of stickers left on the cap and the configured sticker amount in the app differed. It was also mentioned that the colour on the stickers smears.

4.2.4. Improvements

One suggested improvement for the progress feedback would be to include a feature that reminds the app operator to specifically scan the nose and ears at the end of the scan because it is very important that they are included in the scan for a successful calculation of the Cranial Vault Asymmetry Index. Even better would be a feature that detects if the ears and nose are properly captured in the scan.

An improvement for the positioning feedback would be a feature to detect the scanned parts of the head and give instructions on what to scan further and which parts of the head are still missing.

5. Conclusion

In this work, we presented feedback mechanisms of a 3D app that allows physical therapists to create scans of the heads of infants in a real-life setting.

By using information extracted from markers on the cap that an infant is wearing during the scan in combination with depth data captured by the TrueDepth camera system, the app is able to create feedback both about the progress of a scan and the positioning of the iPhone with respect to the infant's head. This feedback is provided to the user by using haptic and audible feedback mechanisms, which allow the user to capture a scan despite the inability to properly see the point cloud on the screen.

In a real-life therapy setting, physical therapists then tested this concept and its implementation. After every scan, they were questioned about the scan experience and its circumstances inside the app and after four weeks of testing, interviews were conducted to further explore the highlights and improvement factors. This showed that the positioning feedback about the distance from the head was very effective at helping all therapists to correct their position. Meanwhile, the positioning instructions for the remaining degrees of freedom were only partially effective, as they were sometimes difficult to realize in practice.

Overall, an easy-to-use interface was achieved by following a coherent design throughout the app. This, however, could only be tested by the physical therapists for the parts of the app that are necessary to create a scan due to their tight schedule. The scan process and user experience were perceived as intuitive and therefore easy to use which was a key objective of this work.

5.1. Future work

Future work should address the cap by designing it so that the markers don't fall off anymore and that the cap is so flexible that each infant can comfortably wear it. Furthermore, the evaluation of the app could be improved by including more infants and scanning operators in the app evaluation. In an effort to make the user more aware of all the features of the app, future work can also include the creation of an onboarding screen.

References

[1] AAP Task Force on Infant, ''Positioning and SIDS,'' Pediatrics, vol. 89, no. 6, pp.1120–1126, Jun. 1992. [Online]. Available: https://pediatrics.aappublications.org/content/89/6/1120

[2] L. C. Argenta, L. R. David, J. A. Wilson, and W. O. Bell, ''An increase in infant cranial deformity with supine sleeping position,'' J. Craniofacial Surg., vol. 7, no. 1, pp. 5–11, Jan. 1996.

[3] K. Ditthakasem and J. C. Kolar, ''Deformational plagiocephaly: A review,'' Pediatric Nursing, vol. 43, no. 2, p. 59, 2017.

[4] E. Ballardini, M. Sisti, N. Basaglia, M. Benedetto, A. Baldan, C. Borgna-Pignatti, and G. Garani, ''Prevalence and characteristics of positional plagiocephaly in healthy full-term infants at 8–12 weeks of life,'' Eur. J. Pediatrics, vol. 177, no. 10, pp. 1547–1554, Oct. 2018.

[5] A. Mawji, A. R. Vollman, J. Hatfield, D. A. McNeil, and R. Sauvé, ''The incidence of positional plagiocephaly: A cohort study,'' Pediatrics, vol. 132, no. 2, pp. 298–304, Aug. 2013.

[6] S. Kluba, F. Roßkopf, W. Kraut, J. P. Peters, B. Calgeer, S. Reinert, and M. Krimmel, ''Malocclusion in the primary dentition in children with and without deformational plagiocephaly,'' Clin. Oral Invest., vol. 20, no. 9, pp. 2395–2401, Dec. 2016.

[7] C.Linz, F.Kunz, H.Böhm, and T.Schweitzer, ''Positional skull deformities: Etiology, prevention, diagnosis, and treatment,'' Deutsches Ärzteblatt Int., vol. 114, nos. 31–32, p. 535, 2017.

[8] R. Stücker, ''Die mit Plagiozephalus assoziierte Säuglingsasymmetrie,'' Zeitschrift Für Orthopädie Unfallchirurgie, vol. 147, no. 4, pp. 503–512, 2009.

[9] M.Feijen, B.Franssen, N.Vincken, and R.R.W.J.van der Hulst, ''Prevalence and consequences of positional plagiocephaly and brachycephaly,'' J. Craniofacial Surg., vol. 26, no. 8, pp. e770–e773, 2015.

[10] X.-Q. Zhao, L.-Y. Wang, C.-M. Zhao, Q. Men, Z.-F. Wu, and Y.-P. Zhang, ''Neurological assessment of Chinese infants with positional plagiocephaly using a Chinese version of the infant neurological international battery (INFANIB),'' Child's Nervous Syst., vol. 33, no. 2, pp. 281–288, Feb. 2017.

[11] Varilag, Plagiocephalus: Ursachen, Spätfolgen & Behandlung von Plagiozephalie. [Online] Available: https://www.varilag.de/ratgeber/plagiocephalus/, accessed 03/15/2022.

[12] Jung, B. K., & Yun, I. S. (2020). "Diagnosis and treatment of positional plagiocephaly". Archives of craniofacial surgery, 21(2), 80–86, https://doi.org/10.7181/acfs.2020.00059.

[13] X. Yang et al., "Mobile3DRecon: Real-time Monocular 3D Reconstruction on a Mobile Phone," in IEEE Transactions on Visualization and Computer Graphics, vol. 26, no. 12, pp. 3446-3456, Dec. 2020, doi: 10.1109/TVCG.2020.3023634.

[14] V. A. Prisacariu, O. Kähler, D. W. Murray and I. D. Reid, "Simultaneous 3D tracking and reconstruction on a mobile phone," 2013 IEEE International Symposium on Mixed and Augmented Reality (ISMAR), 2013, pp. 89-98, doi: 10.1109/ISMAR.2013.6671768.

[15] Barbero-Garcia et al., "Automatic Low-Cost Tool for Head 3D Modelling and Cranial Deformation Analysis in Infants", in Proc. of 3DBODY.TECH 2019 - 10th Int. Conf. and Exh. on 3D Body Scanning and Processing Technologies, Lugano, Switzerland, 22-23 Oct. 2019, pp. 9-14, doi:10.15221/19.009.

[16] P. Tanskanen, K. Kolev, L. Meier, F. Camposeco, O. Saurer and M. Pollefeys, "Live Metric 3D Reconstruction on Mobile Phones," 2013 IEEE International Conference on Computer Vision, 2013, pp. 65-72, doi: 10.1109/ICCV.2013.15.

[17] Kim, Y.M., Mitra, N.J., Huang, Q. and Guibas, L. (2013), Guided Real-Time Scanning of Indoor Objects. Computer Graphics Forum, 32: 177-186. https://doi.org/10.1111/cgf.12225

[18] A. Locher, M. Perdoch, H. Riemenschneider and L. Van Gool, "Mobile phone and cloud — A dream team for 3D reconstruction," 2016 IEEE Winter Conference on Applications of Computer Vision (WACV), 2016, pp. 1-8, doi: 10.1109/WACV.2016.7477629.

[19] OpenCV, OpenCV: Detection of ArUco Markers, Aug. 2022. [Online] Available: https://docs.opencv.org/4.x/d5/dae/tutorial_aruco_detection.html, accessed 08/09/2022.

[20] S. Zeitvogel, C. Wernet, J. Wetzel, A. Laubenheimer and K. Stoevesandt, "RGBD Infant Head Reconstruction for Cranial Vault Asymmetry Estimation," in IEEE Access, vol. 10, pp. 36208-36219, 2022, doi: 10.1109/ACCESS.2022.3160749