

3D MODELLING FOR BENDABLE DEVICES BASED ON INTER-PUPILLARY DISTANCE

Ashish Chopra and *Aman Kumar

Visual Solutions, Samsung Electronics, Noida, India

Received 10th May 2023; Accepted 06th June 2023; Published online 30th July 2023

Abstract

Bendable devices have given a biggest scope in terms of interaction and portability. Besides this several advancement is made to expand its scope for analysis. If we talk about games, motion analysis, medical examination many applications involves capturing person and their movements in 3D. 3D recording characteristically captures the dynamics and movement of the scene during recording and offers the user to change the viewpoint providing the three dimensional model of visualized object. There are conventional 3D recording technologies that provide user with various 3D content. However there is a lack of provision to provide user enriched multimedia content for soothing 3D watching experience. Also no AI model has been yet deployed which can suggest video shooting mode to user. In this paper we have proposed 3D video recording for bendable devices using machine learning. We have aimed to capture a realistic 3D view by predicting the actual depth using bending angle and spectrum analysis. We aim to create a personalized 3D content as per user inter pupillary distance.

Keywords: Inter-pupillary distance, Inter axial distance, Rotating camera, recording modes, Machine learning, Artificial Intelligence

INTRODUCTION

In recent years, various technologies were used to achieve 3D content recording. Conventional 3D recording technologies provides user with various 3D content [1], however personalized 3D content as per user Inter-pupillary Distance (IPD) is still unknown. The assembly consists of a housing to encase two cameras for capturing. In this 3D view has been obtained through two cameras [2]. A stereoscopic 3D camera system comprising: a first camera, which has a first optical axis, a second camera, which has a second optical axis, a beam splitter, a mirror. 3D view has been obtained through two cameras with huge outlet no variable IAD, no rotatable camera and using facial points of that image to match with the face of user captured by the camera sensor and capturing. In this paper, we have proposed 3D video recording for bendable device. A bendable device is used to capture the actual depth using cameras. We have used spectrum analysis technology to analyze the 3D phase and with the use of artificial intelligence shooting mode and bending angles are predicted. In this paper, we are dynamically calculating the inter-axial distance (IAD) between two cameras in a flexible device as per user IPD and displaying user the axis from where user starts bending based on the calculated IAD. The bending of the device leads to changing in IAD of the two cameras, to provide variable depth for 3D recording. Then Correlating object and environment with respect to distance from camera sensor using 3D spectrogram analysis and predict the shooting modes via AI module based on the correlated object and environment. Then, Rotating the camera dynamically in the compulsorily the x-axis in order to record 3D video in the predicted shooting mode. The rotation of the cameras leads to changing convergence which can be used in S3D (Stereoscopic 3D) recording to obtain a personalized 3D video based on user's IPD.

INVENTION TERMINOLOGIES

Inter pupillary distance

Inter-pupillary distance [3], [4] is the distance between the center of pupils. In this paper, we have created a personalized 3D content as per the user's inter-pupillary distance as shown in Fig 1. We have proposed to bend the device for 3D recording based on Inter-pupillary distance.



Fig. 1. The inter-pupillary distance

Inter-axial distance

Fig 2 Shows inter-axial distance which is defined as the distance between the center of the lenses of a stereo camera. We propose to dynamically calculate inter-axial distance and display the user axis from which he starts bending the device.



Fig. 2. The inter-axial distance

*Corresponding Author: Aman Kumar
Visual Solutions, Samsung Electronics, Noida, India

Rotating camera

We have used stereo cameras that is a pair of cameras (Fig 3) that besides providing what a single camera does provides further information like depth information. We have used stereo cameras that are capable of rotating from its axis. This is required to perform shooting in parallel and converged mode.

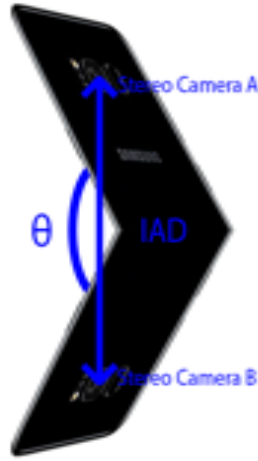


Fig. 3. The pair of cameras

Recording Modes

We propose using of two different recording modes. These are converged and parallel. Fig 4 shows the respective recording modes used.

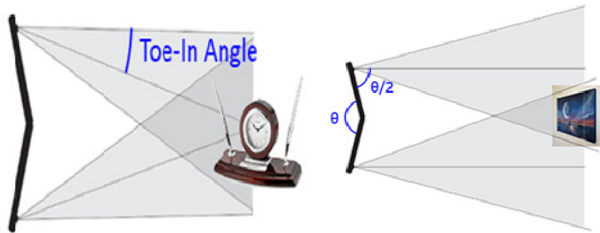


Fig. 4. Converged and parallel recording modes respectively

APPROACH

We have divided the whole process in 4 phases. As illustrated in Fig 5, First phase includes the calculation of IAD based on user's IPD. Fetching of image frames and object distance is also done. In second phase, we calculate the bending angle. Bending angle is calculated as $\Theta = 2\sin^{-1}(IAD/l)$. After calculation of bending angle, in phase three 3D spectrogram analysis is done.

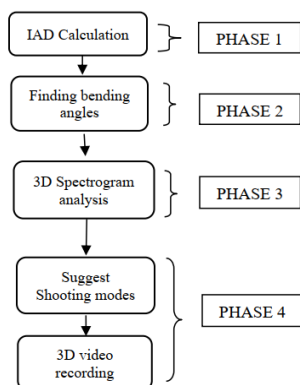


Fig. 5. Process flow

It involves analysis of 3D view of phase and use it for training our algorithm. Later in phase four shooting modes are suggested and based on the suggested shooting modes 3D video recording is done. It is done using various prediction algorithm.

PROPOSED METHODOLOGY

Our invention includes the following stages:

Calculating inter axial distance

The first proposed step comprises of two parts. First we estimate inter pupillary distance (IPD) of user. Next is to fetch image frames and object distance then dynamically calculate inter axial distance (IAD) between two cameras as per user IPD.



Fig. 6. Calculation of inter-axial distance

By setting the Inter-Axial distance to a person's inter-pupillary Distance, we can create customized 3D content so people can relive memories as they experienced it via VR and sound recording. It can also help to make a person experience different locations without being physically present there. As per Military Handbook 743A and the 2012 Anthropometric Survey of US Army Personnel [5], Fig 7 shows mean IPD values for male and female. Using the data from the two cameras, and the known optical zoom level, the person's Inter-pupillary Distance (IPD) can be known.

Gender	Sample size	Mean	Standard deviation	Minimum	Maximum	Percentile				
						1st	5th	50th	95th	99th
Female	1986	61.7	3.6	51.0	74.5	53.5	55.5	62.0	67.5	70.5
Male	4082	64.0	3.4	53.0	77.0	56.0	58.5	64.0	70.0	72.5

Fig. 7. Table showing IPD values (mm) from 2012 Army Survey

To maintain consistency for an individual, we would allow great depth. For far off, big things like an airplane, max IAD is best as it gives best depth. For close off, small things like a coin, IAD has to be set to distance/30 or lower for comfortable viewing. We propose to calculate the distance between the center of the individual's eyeballs by:

1. Compensating for the distance of the two cameras from the eyes while measuring the IPD.
2. Asking the user to look at a very far object (practically looking at infinity) and taking the Inter-pupillary Distance directly.

The choice of the IAD is crucial for creating the required 3D media. There are 3 factors that are needed to be considered while setting the IAD to ensure that there is no divergence.

These are following factors to be considered:

Best mode to record 3D scene

There are two modes to record a 3D scene that appears human realistic:

General Mode (IPD=65mm):

Here **IAD is mean value of IAD of the general population**. If user want to share the 3D video recorded then user will choose general mode, so that it will be not specific to a particular user and give soothing 3D watching experience to every user who is watching 3D video recorded content.

Personalized mode:

Here **IAD is equivalent to user's IAD**. If a user wants to personalize 3D video recorded then user should choose Personal mode, so that it will be specific to that particular user and give soothing 3D watching experience to him. We tend to use the relationship between bending angle and IAD for calculation as $\Theta = 2\sin^{-1}(IAD/l)$.

Dwarfism/Gigantism

Using the data from the two cameras, and the known optical zoom level, a person's Interpupillary Distance (IPD) can be known. However this equalization is not applicable every time. The greater the Inter-Axial Separation, the greater the depth effect. For example as shown in Fig 8, an elephant can perceive much more depth than a human, and a human can perceive more depth than a mouse. However, using this same analogy, the mouse can get close and peer inside the petals of a flower with very good depth perception, and the human will just go "cross-eyed". Decreasing the interaxial will lead to macro stereo photos and separating the cameras will allow great depth.

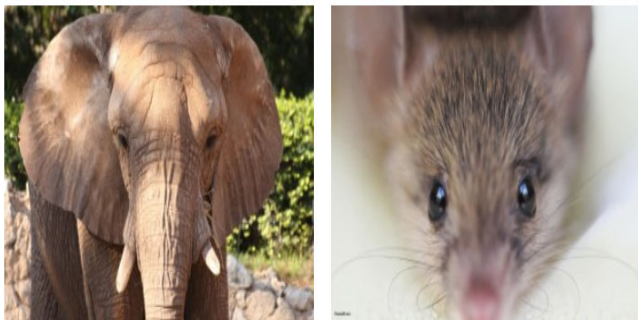


Fig. 8. Images showing hyper-stereo (P.O.V of an elephant) and hypo-stereo effect (P.O.V of a mouse)

1/30 and 1/60 rule

This rule is used for preventing divergence of the eyes (Fig 9) on viewing the 3D content. Divergence never happens naturally, which would mean that your eyes would angle outward. The reason being stereographer [6], [7] should avoid background parallax values in their scene that may require the eyes to diverge when viewed. This is easy to keep in check through some simple math. The 1/30 rule refers to a commonly accepted rule that has been used for decades by hobbyist stereographers around the world. It basically states that inter-axial separation should only be 1/30 of the distance from your

camera to the closest subject. In the case of ortho-stereoscopic shooting that would mean your cameras should only be 2.5" apart and your closest subject should never be any closer than 75 inches (about 6 feet) away. The 1/30 rule gives the maximum IAD as compulsorily distance/30. So, for the subject IAD require should be less than 1/30 rule as mentions above. [(1/30)*distance of object to subject]. The 1/30 rule certainly does not apply to all scenarios. In fact, in feature film production destined for the big screen we will typically use a ratio of 1/60, 1/100 or higher. The 1/30 rule works well if your final display screen size is less than 65 inches wide, your cameras were parallel to each other, and your shots were all taken outside with the background at infinity. Other rules as 1/60 is for comfortable viewing for different screen sizes. Hypo-stereo technique is common for theatrically released films to accommodate the effects of the big screen. It is also used for macro stereoscopic [8] photography. A stereographer calculates the parallax range and uses the various equations available to calculate maximum positive parallax (the parallax of the furthest object), which will translate into a real-world distance when the footage is eventually displayed.

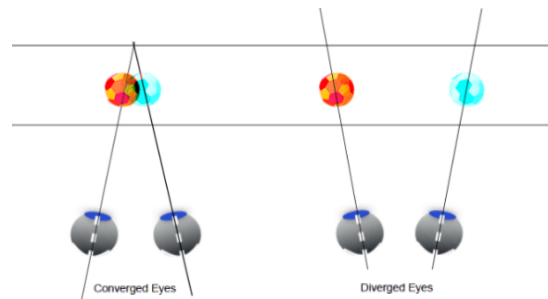


Fig. 9. Images depicting converged and diverged eyes respectively

Finding bending angles

The next step is to calculate the bending angle [9], [10] based on IAD. Suppose distance between the cameras is $D = B_1x+B_2y$ where, $B_1 + B_2 = 1$; $x =$ measurement of first instrument, $y =$ measurement of second instrument.

Standard deviation $\sigma = \sqrt{(B_1^2\sigma_x^2 + B_2^2\sigma_y^2 + 2B_1B_2\sigma_{xy})}$.

Gives $\sigma_D = \sigma_x\sigma_y / \sqrt{(\sigma_x^2 + \sigma_y^2)}$; Where always $\sigma_D \leq \sigma_x$ and $\sigma_D \leq \sigma_y$. Thus we have improved accuracy by reduced Standard Deviation on using another instrument. Hence, as shown in Fig 10, we propose to use the following $\Theta = 2\sin^{-1}(IAD/l)$ where l is the 180 degree distance between cameras which is captured post taking multiple measurements with different IAD's. We display user the axis from user starts bending the device based on calculated IAD. When we bend the device then it leads to change in IAD of two cameras. This provides the variable depth for 3D recording.

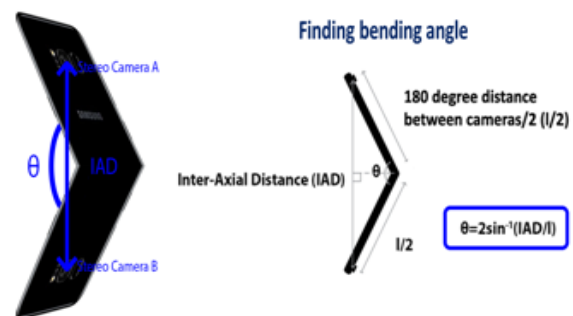


Fig. 10. Calculation of bending angle

The rotation of the cameras [11] in the single x-axis is utilized to enable 3D data capturing on variable IAD. In non-bendable phones the rotatable cameras are not required [12] if we wish to shoot only parallel with lots of image cutting. However, if we bend the phone, the cameras would be incapable of recording 3D data. However in bendable devices, the cameras would rotate compulsorily [13] in the x-axis to enable 3D data generation and enabling parallel and converged shooting modes. Without rotatable camera and a bending device, 3D view is not possible and hence our algorithm is required.

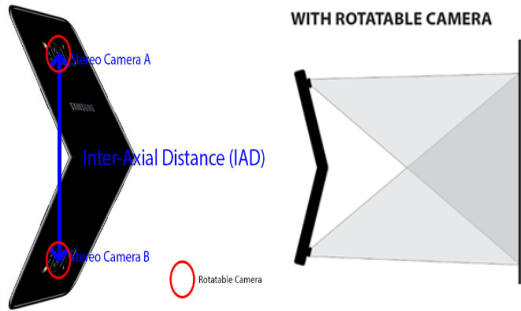


Fig. 11. How rotatable camera along with bendable device is essential for Stereoscopic 3D effect

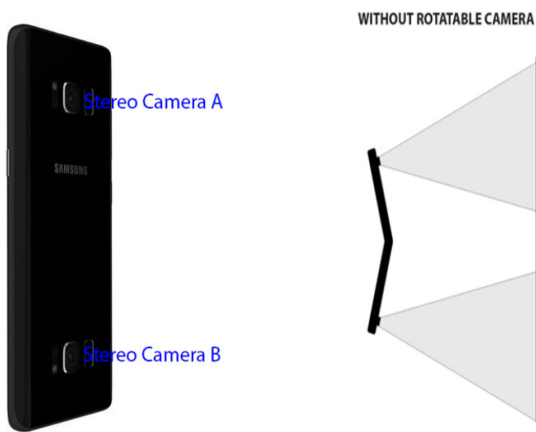


Fig. 12. How non- rotatable camera lacks in 3D view

3D Spectrogram analysis

This step basically involves correlating object and environment with respect to distance from camera sensor using 3D spectrogram analysis. 3D view of the face is analyzed and used for training our algorithm

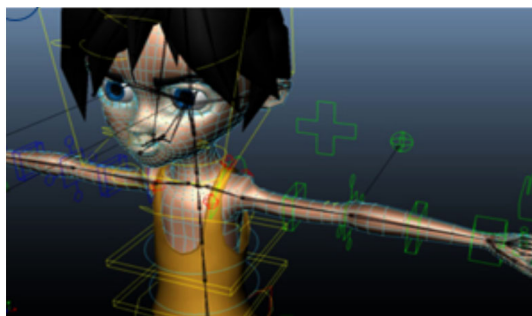


Fig 13. 3D Spectrogram Analysis

Each image is considered to be more encapsulated with more details of view so that in 2D plane [14] we are drawing 3D view through advantages of open GL which is more precise. 3D view image is stored which will be processed and used for

video recording process i.e. using that information to recommend the best shooting mode - parallel or converged, for providing better security by retraining the current model for iris scanner, and for better selfie by reducing the anomalies from the face at run time.

Predicting shooting mode via AI module

The next step will be used to predict the shooting modes via AI module [15] based on the correlated object and environment with respect to distance from camera sensor. So recording would be done on the basis of recommended mode and bending angle. There are two types of shooting modes depending upon distance of object from eyes

1. Shooting converged mode
2. Shooting parallel mode

1. Shooting converged mode: In this, as shown in Fig 15, the point of convergence is at the object of interest (kept at zero parallax). It has an advantage that 3D content is useable without post processing [17], [18]. HIT Cropping is not required so no loss of image dimensions. But the disadvantage is that Keystone effect [19] is observed (Fig 16) and over scan is required to change convergence point later.

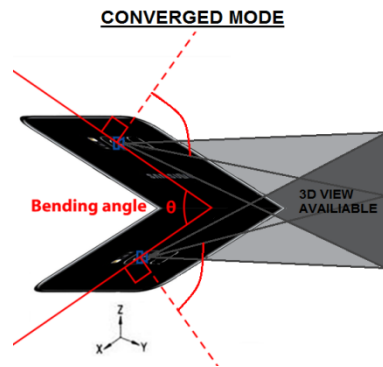


Fig. 14. Converged shooting mode

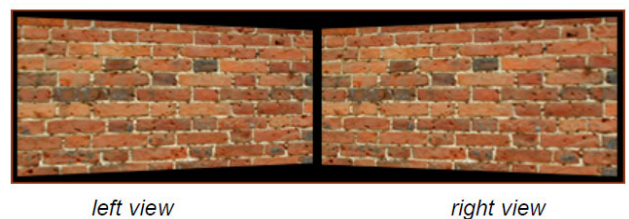


Fig. 15. Keystone effect from converged cameras

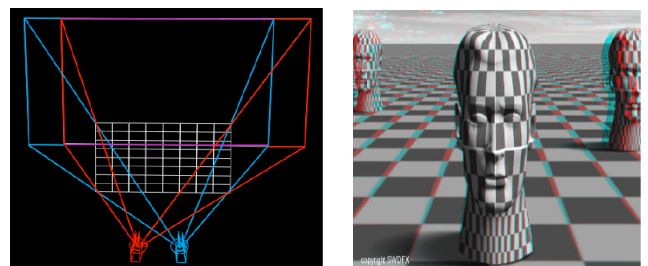


Fig 16 and 17. Converged camera setup and stereo render with converged camera setup

In converged mode, to keep furthest object un-diverged with focus on screen object (f), the following relation is satisfied

$$\frac{\tan^{-1}\left(\frac{IAD}{d}\right)}{\frac{\phi}{2}} < \frac{IPD}{screenwidth} \dots\dots\dots (1)$$

Also, for all objects un-diverged (furthest object at infinity), the following relation is satisfied

$$IAD < 2d * \tan\left(\frac{\phi}{2}\right) * \frac{IPD}{screenwidth} \dots\dots\dots (2)$$

Here, ϕ is the Field of View (FOV) of a single camera and d is the distance between object and camera.

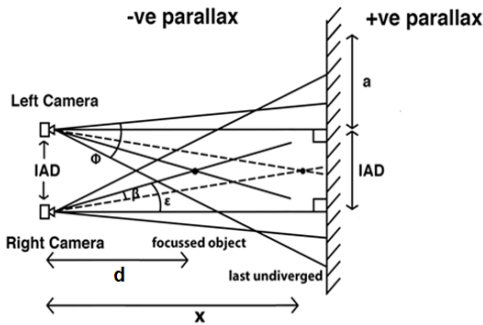


Fig. 19. Converged mode while recording with variable IAD

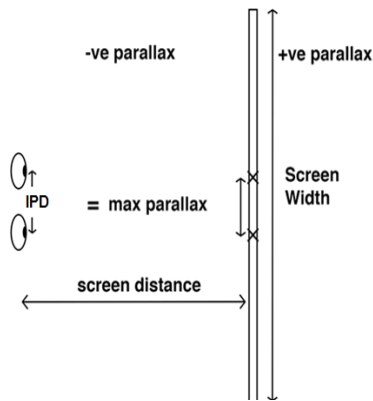


Fig. 20. Actual view equivalent of Fig 19

Shooting Mode parallel

Here, as shown in Fig 21, the point of convergence becomes essentially infinity. It has an advantage over converge mode that no keystone effect is observed. But at the same time Horizontal Image Translation (HIT) fix is required in post processing which leads to a loss in resolution.

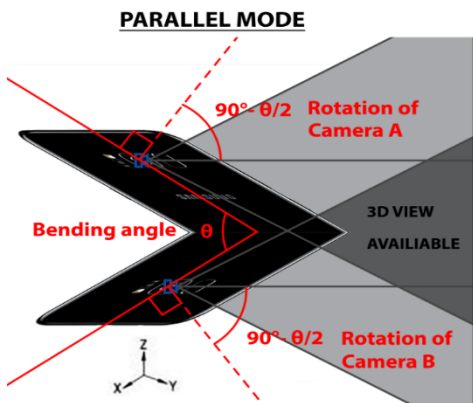


Fig. 21. Parallel shooting mode



Fig 22. HIT has to be applied to merge the two images

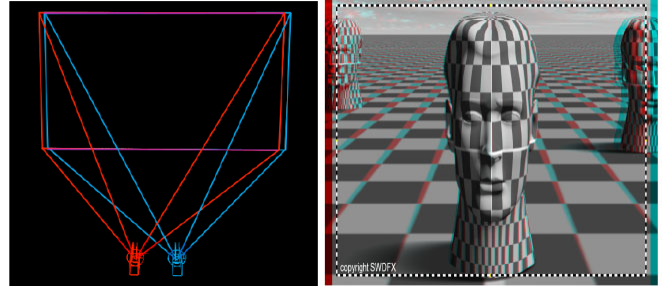


Fig 23 and 24. Parallel camera setup and stereo parallel render with HIT fix

As per Fig 23, the distance between the blue color and red color images should not be greater than IPD. The parallel mode will crop the selected area as shown in Fig 24.

In parallel mode, to keep the furthest object un-diverged with focus on screen object (f), the following relation is satisfied

$$\left(\frac{IPD}{2d * \tan(\phi/2)} - \frac{IPD}{2 * f * \tan(\phi/2) * \tan(\phi/2)}\right) < \frac{IPD}{screenWidth} \dots\dots\dots (3)$$

Also, for all objects un-diverged (furthest object at infinity), the following relation is satisfied

$$IAD < \frac{2d * \tan(\phi/2) * IPD}{screenWidth} \dots\dots\dots (4)$$

Here, ϕ is the Field of View (FOV) of a single camera and d is the distance between object and camera.

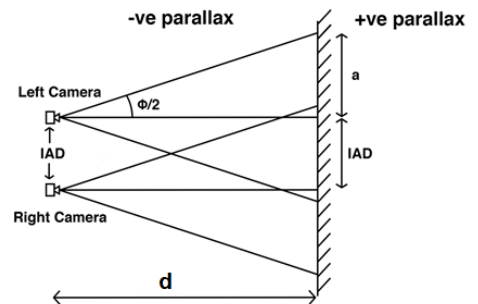


Fig. 25. Parallel mode while recording with variable IAD

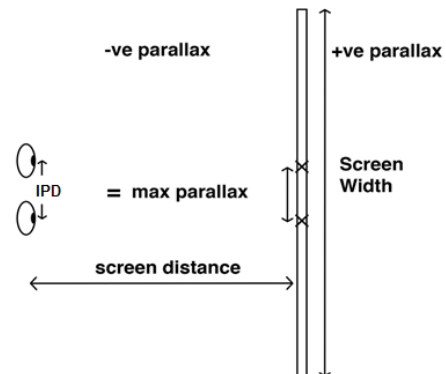


Fig. 26. Actual view equivalent of fig 25

Considering parallel and converged modes we have different angle of tilting the camera. For parallel mode (Fig 27) tilting of camera is given by $90 - \Theta/2$ and for converged mode (Fig 28) we have $\tan^{-1}((IAD/2)/Object\ distance)$.

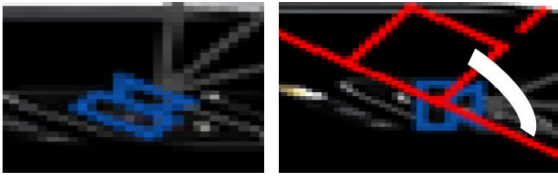


Fig 27. Tilting of camera in parallel mode



Fig 28. Tilting of camera in converge mode

Rotating camera and 3D recording

This involves rotating the camera dynamically compulsorily in the x-axis in order to record 3D video in the predicted shooting mode. The rotation of the cameras leads to change in the convergence which can be used in S3D (Stereoscopic 3D) recording and obtaining personalized 3D view based on the user. Here we would be discussing some specific application pertaining to 3D recording.

Better selfie for user

We can obtain better selfie [20] if we have 3D view of an image. As we have the 3D view of an image we can use it for taking better selfie by inputting more points before clicking pictures and provide best photo to user by removing abnormalities from face.

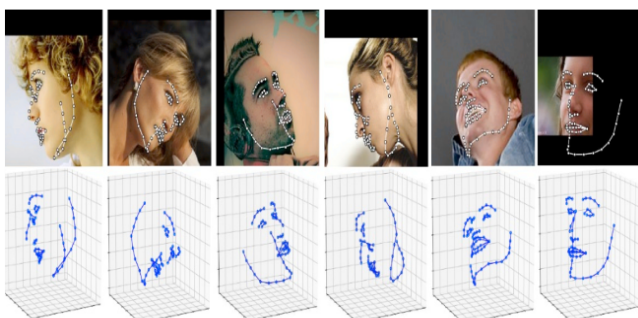


Fig. 28. How 3d view can help predefining abnormalities, spots, depth of eye bag to provide better selfie at runtime

Better Security to current phone unlock system

Our approach can be used in phone unlocking purpose. Suppose phone has two camera where IAD is fixed so every time it will click 2 same images of face to unlock the device.

In conventional 3D image we have unlock function as

$$F(x) = \text{Phone Unlock (Image1, Image 2)}$$

```
# create model
• model = Sequential() [As a instance of Keras model for deep learning and logistic regression for supervised learning]
• model.add(Dense(32, input_shape=(160,), activation='relu'))
• model.add(Dense(16, activation='relu'))
• model.add(Dense(10, activation='relu'))
• model.add(Dense(5,activation = 'softmax'))

# Compile the model
# port in android using tensor flow.
```

Fig. 30. Neural network

But through ours, every time phone uses variable IAD to unlock the phone. So no two images can unlock match 3D model every time as with variable IAD more combinations are there so more security.

Our function is

$$F(x) = \text{Phone Unlock (Image1, Image 2, IAD)}$$

We are calling predict function by passing an image file and returning Boolean yes/no with 100% confidence.

RESULTS

In our paper we have defined two type of shooting modes converged and parallel. Tilt angles are calculated with respect to each mode. To obtain personalised 3D video, a user is advised to bend the mobile phone to that particular tilt angle. While a normal 3D Image through dual camera has depth max =10mm but we are able to reach greater than 50mm depth, so a far better quality of video is presented to the user. More depth will give more soothing effect and feeling realistic 3d view and this is not possible with dual cameras.

Table 1. Recommendations for different shooting modes

Key Parameters for defining mapping function (Parameters => modes)	Shooting Converged	Shooting Parallel
Flatness (Remove Keystone effect)	For regular objects where keystone effect would not be noticeable like telephone	For flat objects like TVs, paintings, brick walls, Theatre screens etc.
Object distance (Removing unnecessary toe in for large distances)	For close objects like flower vase, coffee mug etc.	For far off objects like Sea, beaches, mountains, valleys etc.
Frame size required (Remove HIT cropping)	When we need to have constant resolution or maximum frame size. Use case - regular recording	When the object is far off leading to minimal loss in frame size and when the focal distance is not changing leading to constant resolution.

In generalized mode we will be able to achieve depth more than 50 mm > 10mm because we are including average IPD value.

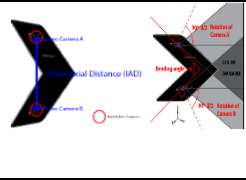



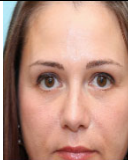
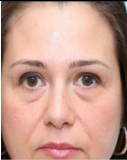
Our Invention (Effectiveness)	This Invention	Conventional
Best mode Predicting shooting modes for best media (image/video) watching experience.		
2 nd mode Provide better safety for unlocking the phone		
3 rd mode Better selfie for user with reducing all anomalies		

Fig. 31. Effects of our invention

Conclusion and Future scope

In this paper we have propose 3D video recording for bendable device. A bendable device is used to capture the actual depth using cameras. We have used spectrum analysis technology to analyze the 3D phase and with the use of artificial intelligence shooting mode and bending angles are predicted. In this paper, we are dynamically calculating the inter-axial distance (IAD) between two cameras in a flexible device as per user IPD and displaying user the axis from where user starts bending based on the calculated IAD. The bending of the device leads to changing in IAD of the two cameras. Thus after undergoing various steps a personalized 3D video is obtained. Our solution provide an enriched multimedia content to the user along with an artificial intelligent solution of shooting mode prediction. In future the research can be extend to projection of multimedia content beyond 3D as 4D, 5D etc. The idea of bendable camera can also enhance the hardware structure of mobile phones.

REFERENCES

- Alpaslan Z. Y. and A. A. Sawchuk, "Multiple camera image acquisition models for multi-view 3D display interaction," in *2004 IEEE 6th Workshop on Multimedia Signal Processing*, 2004, pp. 259–262, doi: 10.1109/mmisp.2004.1436542.
- Peleg S. and M. Ben-Ezra, "Stereo panorama with a single camera," *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, vol. 1, pp. 395–401, 1999, doi: 10.1109/cvpr.1999.786969.
- Dodgson, N. A. "Variation and extrema of human interpupillary distance," in *Stereoscopic Displays and Virtual Reality Systems XI*, May 2004, vol. 5291, pp. 36–46, doi: 10.1117/12.529999.
- Pryor, H. B. "Objective measurement of interpupillary distance," *Pediatrics*, vol. 44, no. 6, 1969.
- Gordon C. C. et al., "2012 Anthropometric survey of u.s. army personnel: methods and summary statistics," Dec. 2014. Accessed: May 20, 2021. [Online]. Available: <https://apps.dtic.mil/sti/citations/ADA611869>.
- Hakala J., M. Nuutinen, and P. Oittinen, "Interestingness of stereoscopic images," in *Stereoscopic Displays and*

Applications XXII, Feb. 2011, vol. 7863, p. 78631S, doi: 10.1117/12.872437.

- Gurrieri L. E. and E. Dubois, "Stereoscopic cameras for the real-time acquisition of panoramic 3D images and videos," in *Stereoscopic Displays and Applications XXIV*, Mar. 2013, vol. 8648, p. 86481W, doi: 10.1117/12.2002129.
- Gao Z., A. Hwang, G. Zhai, and E. Peli, "Correcting geometric distortions in stereoscopic 3D imaging," *PLoS One*, vol. 13, no. 10, p. e0205032, Oct. 2018, doi: 10.1371/journal.pone.0205032.
- Lim S., J. Ha, and D. Lee, "3D Pose and Curvature Estimation of Bendable Interventional Device using Single-view X-ray Image," in *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS*, Jul. 2020, vol. 2020-July, pp. 4732–4736, doi: 10.1109/EMBC44109.2020.9176591.
- Büschel W., P. Reipschläger, and R. Dachsel, "Foldable3D: Interacting with 3D content using dual-display devices," in *Proceedings of the 2016 ACM International Conference on Interactive Surfaces and Spaces: Nature Meets Interactive Surfaces, ISS 2016*, Nov. 2016, pp. 367–372, doi: 10.1145/2992154.2996782.
- Hartley R. I., "Self-calibration from multiple views with a rotating camera," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 1994, vol. 800 LNCS, pp. 471–478, doi: 10.1007/3-540-57956-7_52.
- Nguyen M., H. Le, H. Tran, and W. Yeap, "Adaptive stereo vision system using portable low-cost 3D mini camera lens," in *2017 24th International Conference on Mechatronics and Machine Vision in Practice, M2VIP 2017*, Dec. 2017, vol. 2017-December, pp. 1–6, doi: 10.1109/M2VIP.2017.8211465.
- Shigang L., S. Tsuji, and M. Imai, "Determining of camera rotation from vanishing points of lines on horizontal planes," 1990, pp. 499–502, doi: 10.1109/iccv.1990.139581.
- Chihaoui M., A. Elkefi, W. Bellil, and C. Ben Amar, "A survey of 2D face recognition techniques," *Computers*, vol. 5, no. 4. MDPI AG, p. 21, Dec. 01, 2016, doi: 10.3390/computers5040021.
- "Research Provocation: Understanding the Stereoscopic 3D Cinema Soundscape." https://www.researchgate.net/publication/327848838_Research_Provocation_Understanding_the_Stereoscopic_3D_Cinema_Soundscape (accessed May 20, 2021).
- Eickelberg, S. "Perception of dynamic horizontal image translation in stereo 3D content: An experimental study to identify cognitive factors," in *5th IEEE International Conference on Consumer Electronics - Berlin, ICCE-Berlin 2015*, Jan. 2016, pp. 1–5, doi: 10.1109/ICCE-Berlin.2015.7391204.
- Lee S., D. Jang, E. Kim, S. Hong, and J. H. Han, "A real-time 3D workspace modeling with stereo camera," in *2005 IEEE/RSJ International Conference on Intelligent Robots and Systems, IROS*, 2005, pp. 2140–2147, doi: 10.1109/IROS.2005.1545105.
- Ijsselstein W. A., H. De Ridder, and J. Vliegen, "Subjective evaluation of stereoscopic images: Effects of camera parameters and display duration," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 10, no. 2, pp. 225–233, 2000, doi: 10.1109/76.825722.
- Li B. and I. Sezan, "Automatic keystone correction for smart projectors with embedded camera," in *Proceedings - International Conference on Image Processing, ICIP*, 2004, vol. 4, pp. 2829–2832, doi: 10.1109/ICIP.2004.1421693.
- Schneider T. M. and C. C. Carbon, "Taking the perfect selfie: Investigating the impact of perspective on the perception of higher cognitive variables," *Front. Psychol.*, vol. 8, no. JUN, p. 971, Jun. 2017, doi: 10.3389/fpsyg.2017.00971.