안드로이드기반 스테레오스코픽 3D 기술 특성분석 연구

김정길* 종신회원

A Characterisitc Analysis Study of Android based Stereoscopic 3D Technology

Cheong Ghil Kim* *Lifetime Members*

요 약

최근의 3D 입체 영상 관련 기술의 발달은 스마트폰을 포함한 모바일 단말에서도 3D 입체 영상 서비스의 상용화를 가능하게 하고 있다. 스마트폰에서의 3D 입체 영상 서비스는 무안경 3D 디스플레이와 스테레오스코픽 3D 기술을 이용하고 있으며, 스테레오스코픽 이미지 관련 기술들은 지속적으로 활발히 연구가 진행되고 있다. 3D 입체 영상 서비스를 위하여 각각 다른 표준 H.264/AVC, H.264/AVC SEI, H.264/MVC가 MPEG에서 제정되었다. 본 논문에서는 최근 새롭게 대두되고 있는 모바일 단말에서의 스테레오스코픽 3D 기술의 발전과 특히 안드로이드 폰에서의 3D 서비스를 위한 비디오 포맷과 기능들에 초점에 맞추어 기술적 특성에 관한 분석을 행함으로서 그 활용성을 확인하였다.

Key Words : Stereoscopic 3D; 3D video; H.264/AVC; MVC; smartphone; 3DTV

ABSTRACT

In recent years, the developments in 3D technologies have initiated the commercialization of 3D services on mobile devices. For this purpose, stereoscopic 3D technology is used, which enables 3D TV on mobile devices including smartphone with glasses-free 3D viewing. As a result, the issues related with stereo imaging have been spotlighted greatly. Especially, three MPEG coding standards are provided for mobile 3D services, namely H.264/AVC with and without SEI message and H.264/MVC. In this respect, this paper presents an overview of developments in stereoscopic technologies for mobile devices to gain some perspective on the changes and progress. in this paper, we verified the availability of android based stereoscipic 3D technology related with mobile 3D TV and Smartphone with special emphasis on 3D video format and 3D features by various technollogy characteristics analysis.

I. Introduction

Recent developments on 3D technologies have made 3DTV and mobile 3DTV the most spotlighted technology in the area of audio-video entertainment and multimedia. This development enables auto-stereoscopy or glasses-free 3D viewing on smartphone even without glasses [1-3]. At the same time, there have been many researches on various aspects of 3DTV content creation, coding, delivery, and system integration. As of mobile TV, standardization and legislation activities have lead to creation of similar yet content or country specific standards; for examples, the Korean 3D T-DMB (Terrestrial Digital Multimedia Broadcasting) [4] and Mobile3DTV delivered through DVB-H (Digital Video Broadcasting-Handheld) [5].

Since different types of 3D displays can render 3D video, there are many types of formats to realize 3D video [6]. The principle of 3D video is based on the binocular vision fused by the signals of both eyes. Therefore, the simplest format for 3D video is the stereoscopic video which contains two captured views [7]. Stereoscopic 3D utilizes the human vision system of feeling the depth of the scenes being viewed. It is the ability of our brain to

※ 본 연구는 문화체육관광부 및 한국콘텐츠진흥원의 2013년 콘텐츠산업기술기원사업으로 수행되었음.
*남서울대학교 컴퓨터학과 (cgkim@nsu.ac.kr)
접수일자 : 2013년 5월 20일, 수정완료일자 : 2013년 6월 11일, 최종게재확정일자 : 2013년 6월 19일

fuse together the two images seen by the eyes (the stereo image pair) to form a single image, named the cyclopean image that contains embedded information about depth and an improved resolution of detail [8]. Stereoscopic 3D has its own distinct features, advantages and problems, together with other 3D viewing technologies, holograph and integrated 3D. Therefore, these technologies have been significantly increasing both in research and commercial communities. In the market, the first 3D Android Smartphone, LG Optimus 3D [1], recording, displaying, and shearing glasses-less 3D content, was introduced in 2011.

In this paper, a brief overview of rapid developments in stereoscopic technologies on mobile devices is introduced to gain some perspective on the changes and progress with special emphasis on mobile 3DTV. For this purpose, this paper is organized as followings. Section 2 briefly presents the concept of stereoscopic 3D and overviews the overall structure of 3DTV system. Section 3 briefly describes mobile 3DTV systems: DVB-H and T-DMB. Section 4 concludes the recent status and points to the future research.

Ⅱ. Background

1. Stereoscopic 3D

Stereoscopic 3D system can be devised using two cameras located at two different positions, which may imitate the human visual system known as binocular stereopsis that allows the visual sense to give an immediate perception of depth on the basis of the difference in points of view of the two eyes. It exists in those animals with overlapping optical fields, acting as a range finder for objects within reach. In stereo vision system, the geometry associated with solving this problem is simplified by assuming that the two cameras are coplanar with aligned image coordinate systems.

Fig. 1 shows the basic structure for the stereo image formation and the stereo camera geometry. The center of the lens is called the camera focal center and the axis extending from the focal center is referred to as the focal axis. The line connecting the focal centers is called the baseline, b. The plane passing through an object point and the focal centers is the epipolar plane. The intersection of two image planes with an epipolar plane makes the epipolar line. Let (X, Y, Z) denote the real world coordinates of a point. The point is projected onto two corresponding points, (x_l, y_l) and (x_r, y_r), in the left and right images. The disparity is defined as the difference vector between two points in the stereo images, corresponding to the same point in an object, $v = (x_l - x_r; y_l - y_r)$ [9,10]. In the stereo vision, one of the most difficult areas is to find matching points or features between the left and right images. This is called stereo matching or stereo correspondence problem and has been an intensive area of research using several approaches of matching primitives; area-based, feature-based, and phased-based ones, for decades.

This kind of format can provide the 3D perception but it cannot provide the parallax-adjustable 3D effect [11] suitable for mobile 3D display. Therefore, it might need to have an alternative of stereoscopic video. At the same time, 3D video is mainly aiming at the home application with high-definition (HD) formats which requires high transmission bandwidth. However, in mobile communications, it is inevitable that we have limited bandwidth and heterogeneous networks with different quality. Therefore, for mobile 3D services, suitable 3D video format is important issue.

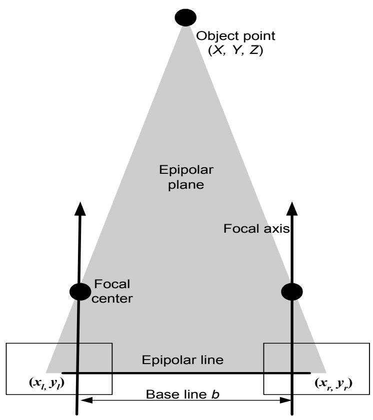

Fig. 1. Basic structure for stereo image formation and stereo camera geometry

2. Mobile 3D Service Systems

Typically, mobile 3D service systems can have several stages from the contents generation to the display on mobile 3D devices. Fig. 2 shows the overall process and the scenario for DVB-H mobile TV can be expressed as follows: stereoscopic video content is captured, properly encoded, encapsulated and then broadcast over mobile TV to be received, decoded and played by a DVB-H enabled portable device.

The first stage is 3D contents generation section with three main approaches: live camera capture, computer generated imagery, and 2D to 3D conversion. For distribution of contents, it may be a way of sending left and right views independently. However, it might be wasteful in terms of bandwidth and packaging for encoding. There are four methods for 3D content encoding: spatial compression, temporal interleaving, 2D plus some form of metadata, and color shifting. There are a number of transmission platforms where 3D content may be deployed: terrestrial broadcast, cable, satellite, packaged material, IPTV, internet download, and mobile TV. In decoding, there are several options and they depend on the encoding, the delivery platform chosen and the display of choice. The following is a list of options as presented by a transcoder manufacturer: external hardware (Set-top boxes, Blu-ray players, DVD players, gaming consoles, decoders), internal hardware (inside the TV or inside the Decoder), firm-ware update to existing devices (Chipsets / STB / Decoders), new hardware (updated chipsets), and software update/download (PC, IPTV). The technology used to simulate the depth presence in a scene influences the type of display used to provide the 3D content. 3D display technologies include: anaglyph, stereoscopy, auto-stereoscopy, holography, and volumetric displays [9].

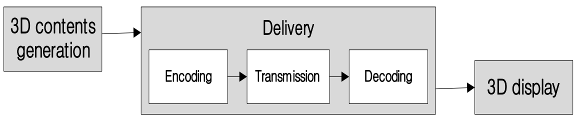

Fig. 2. Overall process of mobile 3D system

Ⅲ. Mobile 3D Systems

1. DVH-T and T-DMD

Mobile 3DTV system was developed by European consortium over DVB-H channel. It naturally consists of different components in which stereo video content is captured, effectively encoded, and then robustly transmitted over DVB-H to be received, decoded and played by a DVB-H enabled handheld [13].

Here, the stereo video framework was adopted while providing comfortable 3D experience to the user with acceptable spatial resolution and frame rate. At the stage of 3D content creation and coding, currently there is no single and generally adopted representation format for stereo video, taking specific mobile channel conditions into account. Most natural is to have two-channel stereo video. Capture of such video by synchronized cameras is relatively easy and the coding can be done efficiently, e.g. by the techniques of the emerging multi-view coding (MVC) amendment of the H.264 AVC standard [14]. Another way for two-channel stereo video is 2D video plus depth format, already standardized under the notion of MPEG-C, Part 3 [14]. This allows good compressibility but requires additional techniques for depth estimation at the content creation side and depth image based rendering at the receiving side [14].

T-DMB, launched in Korea, is MMB (Multimedia Mobile Broadcasting), which delivers multimedia broadcasting services to mobile receivers, handheld receivers, and vehicular receivers even at high speeds matching at least IMT-2000 characteristics. T-DMB data service with video associated data service can provide static or dynamic image, 2D/3D graphics and text data associated to a specific video program using MPEG-4 BIFS, shown in Fig. 3 [15], which is a flexible scene description tool that allows synchronous representation of audio-visual objects in a scene. The BIFS includes information about visual properties of objects for rendering, spatial position of objects and relative time for rendering. Specifically, the proposed 3D T-DMB receiver adopts a look-up table (LUT)-based simultaneous method to accomplish the real time implementation of DIBR algorithms, including warping, hole filling, and interleaving. Moreover, we establish the parameter values that are needed for generating the LUT based on theoretical analysis [15].

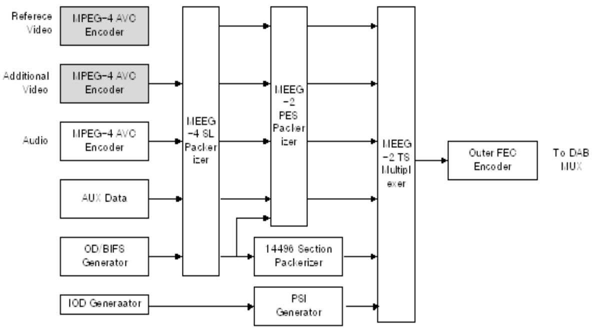

Fig. 3. 3D T-DMB data service based on MPEG-4 BIFS

In addition, the 3rd Generation Partnership Project (3GPP) [16] defined a multimedia container format, known as 3GP (3GPP file format) for 3G UMTS multimedia services and also allowed 3D services. The 3GP file formats is structurally based on the ISO base media file format defined in ISO/IEC 14496-12 - MPEG-4 Part 12; 3GP is similar to MPEG-4 Part 14 (MP4), which is also based on MPEG-4 Part 12. The 3GP and 3G2 file format were designed to decrease storage and bandwidth requirements in order to accommodate mobile phones. Table 1 shows its summary.

Table 1. Comparison between MP4 and 3GP

	MP4	3GP
Spec	MPEG-4 Part 14	MPEG-4 Part 12
Video stream	H.263, H.264/AVC, MPEG-2, MPEG-1	H.263, H.264/AVC, MPEG-4 Part 2
Audio Stream	AAC, MP3, MPEG-4 Part 3, MP2, CELP	AMR-{NB, WB, WB+}, AAC-LC
3D	H.264/AVC MVC	
ETC	MPEG-4 Part 14 includes Part 12 / 3GP, MP4 Scalable.	

2. 3D Smartphone

In accordance with the increasing popularity of smartphone, mobile phone has been quickly evolving from a communications device to multimedia platform. As a result, a significant interest was on focused on stereoscopic 3D services for the development of new technologies and in 2011 the first 3D Smartphone with dual camera system was introduced. It supports 3D video playback as well as records in 3D. It allows the compatibility with YouTube so that users can post 3D home movies directly to YouTube where they can be viewed in 3D right from your handset itself. Also, via an HDMI connection, they can be streamed to any 3D TV.

Table 2 describes various 3D smartphones which were developed on a commercially available mobile hardware platform. This is consisting of four models from LG Optimus 3D [1], HTC EVO 3D [2], Sharp Aquos SH-12C, and Sony Ericsson Xperia Arc [17]. They are Android smartphones and bring visual 3D experiences without glasses. As for application processors, most of them used a dual-core Qualcomm Snapdragon [18] processor except LG Optimus 3D with TI OMAP 4 [19]. Table 3 shows some specification comparisons between major application processors focusing on multimedia and graphics features including 3D codec performance.

Table 2. 3D Android smartphones

	LG Optimus 3D	HTC EVO 3D	Aquos SH-12C	Sony Xperia arc S
System chip	TI OMAP 4430	Qualcomm Snapdragon S3 MSM8660	Qualcomm MSM8255(MSM825)	Qualcomm MSM8255(MSM825)
Processor	Dual core1000 MHz ARM Cortex-A9	Dual core1200 MHzScorpion	Single core 1400 MHz Scorpion	Single core 1400 MHz Scorpion
Graphics processor	PowerVR SGX540	Adreno 220	Adreno 205	Adreno 205
Resolution	480 x 800 pixels	540 x 960 pixels	540 x 960 pixels	480 X 854 pixels
Technology	TFT	S-LCD	LCD	TFT
3D	Glasses-free (Parallax barrier)	Glasses-free (Parallax barrier)	Glasses-free (Parallax barrier)	Glasses-free (Parallax barrier)
Camera	5 megapixels Dual-lens 3D	5 megapixels Dual-lens 3D	8 megapixels Dual-lens 3D	8.1 megapixels Dual-lens 3D
Camcorder	1920x1080 (1080p HD) (24fps) 1280x720 (720p HD) (30fps)	1280x720 (720pHD)	1280x720 (720pHD)	1280x720 (720pHD)
3D capture	Photos: 3 megapixels Videos: 1280x720 (720p) (30fps)	Photos: 2 megapixels Videos: 1280x720 (720p)	Photos: 2 megapixels Videos: 1280x720 (720p)	Photos: 2 megapixels Videos: 1280x720 (720p)

3. OMAP 4 Platforms

This section introduces TI OMAP 4430 [19] platform by Texas Instruments which is used for LG Optimus 3D for stereo-video decoding and playing. Specifically, the OMAP 4430 integrates PowerVR SGX5x cores for superior 2D and 3D graphics rendering by Imagination Technologies [20]. Furthermore, Video capture and playback are both accelerated by the platform's IVA-HD hardware engine. Fig. 3 and Fig. 4 show the block diagrams of stereoscopic recording and playback, respectively.

The procedure of stereoscopic recording is as follows:

- Dual cameras capture left and right views of the scene
- Camera mechanical misalignment correction and autoconvergence applied on left and right views
- H.264 encoder with SEI header for efficient management and storage of 3D content
- Live view of recording is rendered in 3D on selected display terminal

The stereoscopic playback follows the bellowing procedure:

-H.264 SEI decoder reconstructs the 3D packed frames from file system or remote streaming
-JPEG encoder reconstructs the 3D packed image from file system or remote streaming
-Display Subsystem renders 3D content based on the rendering option on selected display terminal

Table 3 summarized the stereoscopic support OMAP4430 features combining with hardware components.

Table 3. Applications processors

Application Processor	MSM8660	OMAP4	MSM825	Rockchip 2918
CPU	Scorpion	Dual Cortex A9	Scorpion	Cortex A8 core
CPU Clock	1.2GHz	1GHz	1.4GHz	1.2GHz
Codec DSP	QDSP-400 Mhz	IVA-HD - 306Mhz	QDSP-256 Mhz	IVA-HD - 306Mhz
Supported Codecs	MPEG-4,MPEG-2,H.264,H.263,VC-1,DivX	MPEG-4, MPEG-2, H.264,H.263,VC-1, DivX	MPEG-4,MPEG-2,H.264,H.263,VC-1, DivX	H.264, VP8, RV, WMV, AVS, H.263, MPEG4. etc
Encoder Performance (2D)	Full-HD@30fps (Baseline profile + B-frames)	Full-HD@30fps (Baseline profile + B-frames)	HD@30fps (Baseline profile + B-frames)	Full-HD@30fps (Baseline profile + B-frames)
Decoder Performance (2D)	Full-HD@30fps	Full-HD@30fps	HD@30fps	Full-HD@30fps
3D Codec Support	N/A	Support	N/A	Support
Possible 3D Codec Performance	720P@30fps (Side-by-Side Based)			

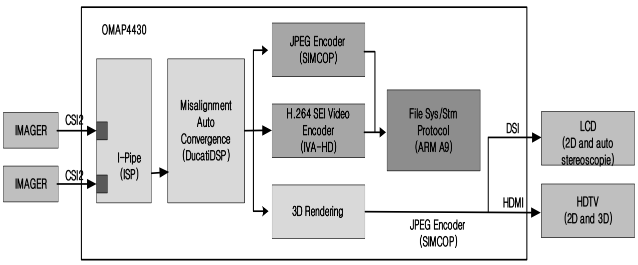

Fig. 3. Stereoscopic recording

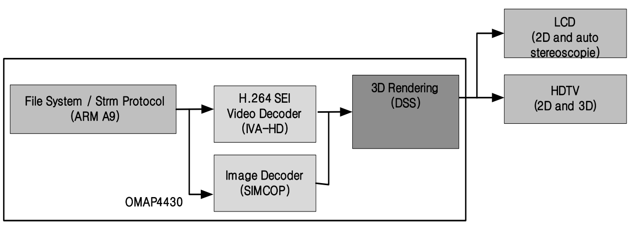

Fig. 4. Stereoscopic playback

Ⅳ. Conclusion

In this paper, we briefly verifiedd the current thenology stage of stereoscopic 3D technologies on mobile with examples of mobile 3DTV systems and 3D smartphones. Due to the rapid technical evolution on mobile industry, the optimal platform for 3D services is evolving continuously with various acceleration features both in software and hardware. Therefore, there will be more possibilities to expand stereoscopic 3D mobile application. Future work will cover more detailed works considering power-constrained mobile platform.

References

[1] LG Optimus 3D smartphone http://en.wikipedia.org/wiki/LG_Optimus_3D.
[2] HTC Evo 3D smartphone http://en.wikipedia.org/ wiki/ HTC_Evo_3D
[3] Sharp Aquos Phone SH-12C http://www.android.com/devices/detail/sharp-aquos-phone-sh-12c
[4] H. Lee, S. Cho, K. Yun, N. Hur, J. Kim, “A Backward-compatible, Mobile, Personalized 3DTV Broadcasting System Based on T-DMB, Three-Dimensional Television Signals and Communication Technology, Springer, pp. 11-28, 2008.
[5] Mobile 3DTV website http://sp.cs.tut.fi/mobile3dtv/
[6] A. Vetro, A.M.Tourapis, K. M¨uller, T. Chen, “3D-TV Content Storage and Transmission”, IEEE Transactions on Broadcasting, Vol. 57 (2), pp. 384-394, June 2011.
[7] Song Ci, Hui Tang, and Yun Ye, “A transcoding framework with error-resilient video/depth rate allocation for mobile 3D video streaming”, IEEE International Conference on Communications (ICC), pp. 2038-2042, 2012
[8] E.A. Edirisinghe and J. Jiang, Stereo imaging, An emerging technology,” Proc. of SSGRR 2000: International Conference on Advances in Infrastructure for E-Business, Science, and Education on the Internet, pp. 1-10, Aug. 2000.
[9] C.G. Kim, V.P. Sirni,and S.D. Kim, “High Performance Coprocessor Architecture for Real Time Dense Disparity Map,” Journal of Korea Information Processing Society, vol 14-A(5), pp. 301-308, 2007.
[10] L. Onural, T. Sikora, J. Ostermann, A. Smolic, R. Civanlar, and J. Watson, "An assessment of 3DTV technologies," Proc. NAB 2006, Las Vegas, NV, pp. 456 - 467, Apr. 2006.
[11] C. Fehn, “A 3D-TV system based on video plus depth information”, Proc. of the 37th Asilomar Conference on Signals, Systems and Computers, vol. 2, pp. 1529-1533, Nov. 2003.
[12] R. Piroddi,"Stereoscopic 3D Technologies," White Paper Snell Ltd., http://www.snellgroup.com.

[13] A. Gotchev, A. Smolic, S. Jumisko Pyykk, D. Strohmeier, G. B. Akar, P. Merkle, and N. Daskalov,"Mobile 3D television: Development of core technological elements and user centered evaluation methods toward an optimized system," Proc. SPIE 7256, Multimedia on Mobile Devices 2009, vol. 7256, pp.1-6, Jan. 2009.

[14] A. Smolic, K. Mueller, N. Stefanoski, J. Ostermann, G. Gotchev, G.B. Akar, G. Triantafyllidis, A. Koz, "Coding Algorithms for 3DTV-A Survey,"IEEE Transactions on Circuit Systems Video Technology, 17 (11), pp. 1606-1621, Nov. 2007.

[15] K. Yun, H. Lee, N. Hur,and J. Kim, "Development of 3D video and 3D data services for T DMB, Stereoscopic Displays and Applications XIX" Proc. of the SPIE, vol. 6803, pp. 68030Z-68030Z, 12, 2008.

[16] The 3rd Generation Partnership Project http://en.wikipedia.org/wiki/3GP_and_3G2

[17] Sony Ericsson Xperia Arc http://en.wikipedia.org/wiki/Sony_Ericsson_Xperia_Arc

[18] Qualcomm Snapdragon Processors http://www.qualcomm.com/snapdragon

[19] Texas Instrument OMAP 4 Series http://www.ti.com/lsds/ti/omap-applications-processors/omap-4-processors-products.page

[20] Imagination Technology PowerVR SGX5x http://en.wikipedia.org/wiki/PowerVR

저자

김 정 길(Cheong Ghil Kim) **정회원**

- 1987년 : Univ. of Redlands, U.S.A. 컴퓨터과학 학사졸업
- 2003년 : 연세대학교 컴퓨터과학 공학석사
- 2006년 : 연세대학교 컴퓨터과학 공학박사
- 2006년~2007년 : 연세대학교 컴퓨터과학과 박사후연구원
- 2007년~2008년 : 연세대학교 컴퓨터과학과 연구교수
- 2008년~현재 : 남서울대학교 컴퓨터학과 교수

<관심분야> : 멀티미디어 임베디드 시스템, 이기종 컴퓨팅, 모바일 AR, 3D Contents