

Computer Facial Animation: A Review

Heng Yu Ping, Lili Nurliyana Abdullah, Puteri Suhaiza Sulaiman, and Alfian Abdul Halin

Abstract—Computer facial animation is not a new endeavour as it had been introduced since 1970s. However, animating human face still presents interesting challenges because of its familiarity as the face is the part used to recognize individuals. Facial modelling and facial animation are important in developing realistic computer facial animation. Both modelling and animation is dependent to drive the animation. This paper reviews several geometric-based modelling (shape interpolation, parameterization and muscle-based animation) and data-driven animation (image-based techniques speech-driven techniques and performance-driven animation) techniques used in computer graphics and vision for facial animation. The main concepts and problems for each technique are highlighted in the paper.

Index Terms—Computer graphics, data-driven animation, facial animation, geometric-based modelling.

I. INTRODUCTION

Computer facial animation is experiencing a continuous and rapid growth since the pioneering work of Frederick I. Parke [1] in 1972. This is partly due to the increasing demand of virtual characters or avatars in the field of gaming, filming, human computer interaction and human machine communication.

The human face is one of the channels in expressing the affective states. It has complex but flexible three-dimensional (3D) surfaces. To present the human face onto a computer system is a challenging task due to several goals to be achieved. According to Deng and Noh [2], the main goal in facial modelling and animation research is to develop a high adaptability system that creates realistic animation in real time mode where decreasing manual handling process. Here, *high adaptability* refers to the system that is easily adapted to any individuals' face.

Various approaches have been proposed by the research community to improve the performance of facial animation in different aspects. Some claims that a good facial modelling and animation involving how to lip synchronization and how to use the eyes, brows and lids for expression [3]. Others believed that a facial model should include other attributes such as surface colours and textures [4].

The focus of this paper is to present a review of the past and recent advancements in facial animation research. Besides, an overview of the facial animation framework is discussed to give a picture of where the modelling and animation techniques take place. The goal of this review is to improve the current computer facial animation system with

the knowledge of the current advancement in this technology.

II. A GENERAL FACIAL ANIMATION FRAMEWORK

Animating facial action is a tedious and complicated task. The construction of facial animation proceeds in several stages and is shown in Fig. 1.

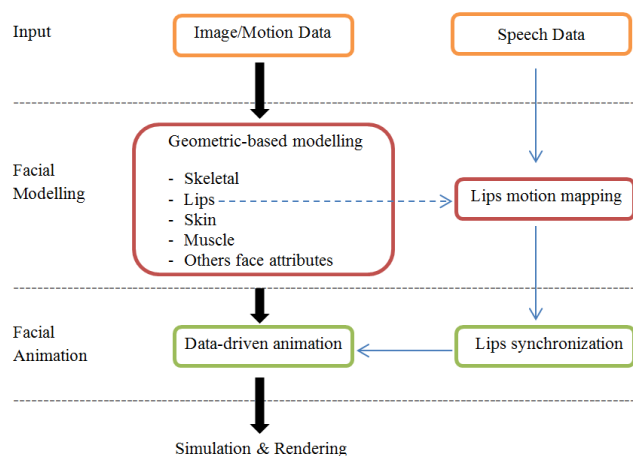


Fig. 1. Facial animation framework

Each facial animation starts from the head model creation. The facial geometry is created from the input data either captured from images or motions. The input data are pre-processed and filtered, only the corresponding data is used to create the basis of morphable meshes. Next, geometric-based modelling is applied to create different parts and details of the face such as the skull, skin and others. At this stage, different face attributes require different modelling techniques. Muscle-based modelling may best used to model the facial muscle but not for the skeletal modelling. Details of the facial modelling techniques are discussed in Section IV. After modelling the face, some additional processing such as texture mapping is applied before facial animation techniques (Section V).

To develop a realistic human facial animation, other facial attributes animation can be adapted to improve the performance. For example, speech-driven animation (Fig. 1, right) where the lip motion is mapped with the lip model created in geometric-based modelling is then being synchronized with facial animation. Finally, the facial animation is followed by the simulation and rendering process before presenting the final output to the users.

III. FACIAL REPRESENTATION

A facial model consists of different parts and details. To give a better insight in modelling and animating faces, the study of anatomy and physiology of the face is needed. Generally, a human-like facial model is a sum of skin,

Manuscript received October 10, 2012; revised January 15, 2013.

The authors are with the Faculty of Computer Science and Information Technology, Universiti Putra Malaysia, Malaysia (e-mail: {hyuping, liyana, suhaiza, alfian}@fsktm.upm.edu.my).

skeletal anatomy, muscular anatomy, and vascular system.

A. Skin

Skin is the external surface of the human forms and comprises by epidermis, dermis and hypodermis. The skin textures may vary due to age, sex, races, thickness, environmental and disease. Texture attributes such as wrinkles is different for a healthy lifestyle female aged 18 compared to a female aged 45 with smoking habit. Thus, facial model should be modelled carefully to reflect the actual physical characteristics [4]. However, there is a trade-off between modelling real physical skin and rendering speed in real time interactive facial model system.

B. Skeletal Anatomy

Skeletal or in other words skull is important for facial synthesis as it provides the framework to place the muscles and skin [4]. The deformation of the skull will determine the shapes of the face. Skull consists of two major components which is the cranium and mandibles [5]. The cranium is in which the brain is lodged whereas the mandibles is the only mobile bone of the facial skeleton on the lower face part. Again, there are differences between male skull and female skull in different age group, races and geographically distant locations. To date, most of the current facial models are either based on female skull or male skull with low adaptability. An African female adult's skull is hardly to adapt into an Asian male adult. Hence, to develop a facial model that meets all the skull considerations with high adaptability is difficult and hardly impossible.

C. Muscular Anatomy

The facial muscles lie between the bone and the skin which provides the physiological functions to drive the facial expression. There are different type facial muscles to determine the facial movement. The eye muscles use to control the eye movement such as opening and closing of the eyelids. The jaw muscles control jaw and tongue position when the facial muscles are performing some chewing or swallowing functions [4]. Facial muscles appear in various shapes, to accurately model and simulate all the muscles is a challenging task. The importance of muscular anatomy in facial model development had led a trend in muscle-based modelling.

D. Vascular Systems

Recently, simulating vascular system that supplies the blood to the facial muscles has become a focus in current facial animation research. The facial skin may blush as the blood supply increased in the arteries whereas react pallor during vice versa condition. These had given the modelled face a 'life' to react more realistically.

Besides the four facial representations highlighted above, there are others facial features such as eye, lip, teeth and tongue which contribute to a real, human-like computer facial animation. Further details of the facial anatomy can refer to the Gray's Anatomy [6].

IV. GEOMETRIC-BASED FACIAL MODELLING TECHNIQUES

In computer graphics and vision, faces represent in

two-dimensional (2D) or three-dimensional (3D) geometry. Geometric-based techniques usually deal with the surface generation, manipulation and deformation in the geometry processing pipeline. Facial modelling is important as the structure of facial model determine its animation potentials [4]. In this section, we will discuss on geometric-based facial modelling techniques which include shape interpolation, parameterization, and muscle-based modelling.

A. Shape Interpolation or Blendshapes

Shape interpolation or blendshapes have been widely used in facial animation research. The fundamental of this method is that during the animation, the interpolated facial model is created from the key frame before and after through interpolation over a normalized time interval (Fig. 2) [1]. This process involves modelling polygonal meshes of a human face to approximate expressions and then blending the different sub meshes automatically [7].

Shape interpolation can be extended in different ways. Bilinear interpolation can be used to generate more possibilities of facial expression changes [8]. Spline interpolation can be applied to create a faster animation effect at the beginning and slower effect at the end of animation [9].

Shape interpolation is mostly applied in the animation which involved simple transformation such as scaling or rotating. Although these interpolations are easily computed, it is restricted to create a wide range of realistic facial animation due to the complex construction of the face. To get realistic facial animation effects, animators have to manually handle each facial model in order to get the perfect blendshape which is time consuming and somehow labour intensive. Therefore, shape interpolation is practical for animation which involve few key frames during early stage. Lately, more and more techniques have been derived from this shape interpolation concept such as pose space animation [10], orthogonal blendshapes [11] and direct-manipulation blendshapes [12].

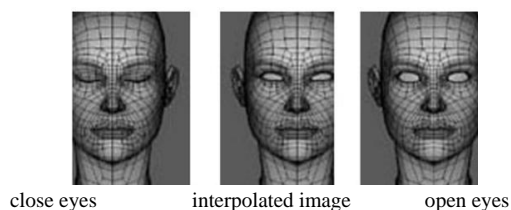


Fig. 2. Linear interpolation is performed on eyes from close to open.

B. Parameterization

To overcome the limitations of shape interpolations, parameterization is being proposed where the basis principle of shape interpolation still remains: representing the face in polygonal meshes and utilizing the key frame interpolation [13]. In parameterization, facial geometry is broken into parts where each part is exposed properly to its parameter sets or control points. This allows the animators to have control of the facial configurations. By combining different parameters, a large range of facial expressions can be produced. Action Units (AU) (Table I) or MPEG-4 Facial Animation Parameters (FAPs) are established to parametrically represent the facial animations in order to smooth the parameterization process.

TABLE I: SAMPLE OF SINGLE FACIAL ACTION CODING SYSTEMS (FACS)

AU	FACS Name	AU	FACS Name
1	Inner Brow Raiser	7	Lid Tightener
2	Outer Brow Raiser	9	Nose Wrinkles
4	Brow Lower	10	Upper Lid Raiser
5	Upper Lid Raiser	12	Lid Corner Puller
6	Check Raiser	14	Dimpler

To parameterize the facial model in parts is very depends on the facial mesh topology. For those simple facial models like CANDICE-3 [14], facial animation can be parameterized relatively intuitive as it just above one hundred vertices [15]. However, for a complex facial model which yields a very large number of vertices, parameterization is not practical. Furthermore, to adapt the existing parameterization into a new facial model require manual tuning process which is tedious and complex [2]. To resolve some of these problems, deformation based technology, multidimensional parameter space techniques [16] and partial differential equations [17] are developed to parameterise facial animation.

C. Muscle-Based Modelling

Platt and Badler [18] pioneered muscle-based model since 1980s. They proposed a mass-spring model to connect the skin, muscle and bone nodes. In Platt's model, spring network is used to interconnect the 38 regional muscles with action units. Late 1980s, Waters [19] proposed a more developed vector muscle model which is the basis principles for today's physic-based model. He classified muscles into three types: linear (Fig. 3), sheet and sphincter (Fig. 3); based on the nature of the muscle action. He also used the FACS to relate the facial expressions to muscle activity.

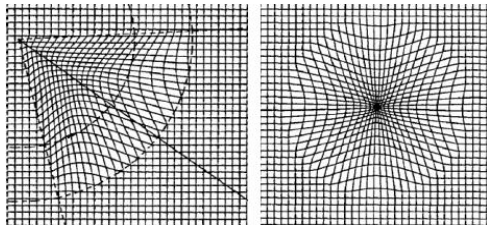


Fig. 3. Waters's muscle model. Linear muscle (left), sphincter muscle (right)

In 1990s, Terzopoulos and Waters [20] introduced a facial model that incorporates both anatomically-based muscle model and physically-based tissue model. The three layer spring mesh are skin, fatty tissues and muscles which tied to bone aimed to model more realistic anatomical facial behaviour. Their work is later improved by others researcher with the introducing of others physics-based muscle model [21], [22] and NURBS muscle based system [23], [24]. Today, muscle-based modelling still received attention from the research community as there are some potential problems to deal with. Among them includes to accurately generate muscle model during different activities in real time mode, to feed in more details of muscle to improve the realism of the facial model and synthesizing the existing muscle model with the data-driven method.

V. DATA-DRIVEN FACIAL ANIMATION TECHNIQUES

A static facial model with just simple blending is no longer enough to support the computer animation pipelines. To facilitate this problem, a number of data-driven techniques

have been proposed. Data-driven techniques widely used as it presents richer, more detailed and more accurate animation from the empirical data collected. We review some present data-driven techniques used in facial animation in the following section: image-based techniques, speech-driven techniques and performance-driven animation.

A. Image-Based Techniques

Image-based techniques are developed to produce a photorealistic human face model that would not be done by only adapting shape interpolations and muscle deformer. In this technique, facial surface and position data is captured from images from different view to reconstruct the facial model. The depth of the model can be computed by applying triangulation on two images. Noh and Neumann [25], grouped the image-based techniques into four which include morphing between photographic images, texture manipulation, image blending and vascular expression. The number and view of input images and how much is known about the scene geometry determined the effectiveness of the system.

In 2000s, Borshukov et al. [26] have demonstrated a sophisticated image-based technique for the movie "The Matrix Reloaded". Yet, image-based techniques were used as special effects in some recent movies such as Star Wars and The Lord of the Rings [27]. On the other hand, for modelling research purpose, image-based techniques is combined with others geometric-based techniques to improve the facial model deformation [27], [28].

Image-based techniques still popular used by animators and researchers as it can produce realistic-looking facial models with low costs compared to 3D laser scan facial model. To remain its popularity, current image-based techniques constraint such as how to automate represents the 3D facial model by using a few parameters during animation should be treated.

B. Speech-Driven Techniques

Another challenging topic in facial animation is to synthesize realistic visual speech animation corresponding to input speech. Previous research efforts in speech-driven animation start from mapping the phoneme annotation to the animation either manually [29] or automatically [30] to produce desire mouth shapes. *Coarticulation* is the phenomenon that is important in lip-synchronization. It stated that the motion of the lip depends on the phoneme currently pronounce, and also those coming before and after [31].

In speech-driven techniques, a pre-recorded facial database is needed for synthesis purposes [2]. *Video Rewrite* by Bregler et al. [32] is a typical speech-driven techniques example where the triphones from the constructed databases is concatenated to create a new audiovisual sequence. Training algorithms such as Hidden-Markov and Gaussian mixture always play important roles to improve the efficiency of mapping voice to face from the databases.

To engage the emotions to the existing speech-driven model, Y.Cao et al. [31] proposed an expressive speech-driven facial animation where the accuracy of the lip-synchronization is presentable. Besides, to develop a high

performance real time lip-synchronize facial animation, Y.M. Chen et.al [33] presented a Dominated Animeme Models (DAM) which integrate the latent dominance function and intrinsic animeme function in order to precisely fit the coarticulation.

There is still room for improvement to synthesize lip-synchronization in real time and incorporate emotions in speech-driven facial animation: How to precisely synchronize the lip motion to deal with real time facial animation which involves dragging sounds? How to identify and represent the best expressive speech during animation played?

C. Performance-Driven Animation

Performance-driven animation involves using motion data captured by some input devices such as data gloves and motion capture cameras to drive the synthetic character. It is a popular technique used in computer animation industry [4] as the animators can capture movements directly from real faces and transfer the data to digital forms.

Performance-driven technique was first introduced by Lance Williams [34], which used a few markers on the face and video camera to scan the face and track the markers. Guenter et.al [35] is then used the technique to derive the data from video streams. Kouaudio et.al [36] developed a real time facial animation system from pre-modelled facial expressions to animate synthetic characters. They introduced markers associated to points where the Euclidean distance between the corresponding points and markers is being minimized to create the intermediated facial expression.

There are many technical challenges in developing a real time performance-driven facial animation which includes how to precisely track the rigid and non-rigid motion of the user's face, and how to map extracted tracking parameters that drive the facial animation [37]. To resolve the above problems, Thibaut Weise et.al [37] has demonstrated a markerless acquisition system that use a face tracking system which is a product of geometry and texture registration.

In order to increase the realism of the generated facial model, performance-driven facial animation can integrate with others techniques discussed above or others detail effect like hairs and neck movement with fewer markers involved. Besides, an intelligent performance-driven facial animation which can adapt input motion data corresponding to the current facial animation requirement will certainly benefit the animators and researchers.

VI. CONCLUSION

Developing a facial animation involves determining relevant geometric descriptions to represent the facial model. Yet, the constructed facial model should be able to support the animation. In this paper, we describe the facial animation framework and some anatomic considerations in representing a face in digital form. Besides, we discuss and review several techniques used in driving the facial animation: geometric-based facial modelling and data-driven facial animation. Within each technique, we describe the main ideas and compared their strength and weakness. Most techniques are actually synthesized by several methods in

order to achieve better performance. Based on the review, we highlighted some potential problems which remain worthy to be explored.

REFERENCES

- [1] F. I. Parke, "Computer generated animation of faces," in *Proc. the ACM Annual Conference*, vol. 1, pp. 451-457, 1972.
- [2] Z. Deng and J. Noh, "Computer facial animation: a survey," in *Data-Driven 3D Facial Animation*, London: Springer-Verlag, ch. 1, pp. 1-28, 2008.
- [3] J. Osipa, *Stop Staring: Facial Modelling and Animation Done Right*, Indianapolis, Indiana: Wiley Publishing Inc, pp. 3-30, 2007.
- [4] F. I. Parke and K. Waters, *Computer Facial Animation*, Wellesley, MA: AK Peters Ltd, pp. 85-362, 2008.
- [5] S. Garchery, T. D. Giacomo, and N. Magnenat, "Real-time adaptive facial animation," in *Data-Driven 3D Facial Animation*; London: Springer-Verlag, ch. 12, pp. 217-247, 2008.
- [6] S. Standring, *Gray's Anatomy*, 40th ed. New York: Churchill-Livingstone-Elsevier, 2009.
- [7] C. Liu, "An analysis of the current and future state of 3d facial animation techniques and systems," M. S. thesis, School of Interactive Arts and Tech., Simon Fraser Univ., Burnaby, BC, Canada, 2009.
- [8] K. Arai, T. Kurihara, and K. Anjyo, "Bilinear interpolation for facial expression and metamorphosis in real time animation," *The Visual Computer*, pp. 105-116, 1996.
- [9] K. Waters and T. Levergood, "DECface: An automatic lip-synchronization algorithm for synthetic faces," *Technical Report 93/4, Digital Equipment Corp*, Cambridge Research Laboratory, 1993.
- [10] B. Bickel, M. Lang, M. Botsch, and M. Gross, "Pose-space animation and transfer of facial details," in *Proc. the 2008 ACM Siggraph/Eurographics Symposium on Computer Animation*, pp. 57-66, 2008.
- [11] Q. Li and Z. Deng, "Orthogonal-blendshape-based editing system for facial motion capture data," *IEEE Computer Graphics and Applications*, pp. 76-82, Nov./Dec. 2008.
- [12] J. P. Lewis and K. Anjyo, "Direct manipulation blendshapes," *IEEE Computer Graphics and Applications*, pp. 42-50, July/Aug. 2010.
- [13] M. Radovan and L. Pretorius, "Facial animation in a nutshell: past, present and future," in *Proc. SAICSIT 2006*, pp. 71-79.
- [14] J. Ahlberg, "CANDICE-3: an updated parameterized face," *Technical Report No. LiTH-ISY-R-2326*, Linkoping University, Sweden, January 2001.
- [15] Y. Sheng, A. Sadka, and A. Kondo, "Automatic single view-based 3d face synthesis for unsupervised multimedia applications," *IEEE Trans on Circuits and Systems for Video Technology*, pp. 961-974, 2008.
- [16] A. Arya and S. D'Paola, "Multi-space behavioral model for face-based affective social agents," *Journal of Image and Video Processing, Special Issue on Facial Image Processing*, vol. 2007, pp. 12, 2007.
- [17] Y. Sheng, P. Willis, G. Castro, and H. Ugail, "PDE-based facial animation: making the complex simple," *International Symposium on Visual Computing 2008*, part II, LNCS, pp. 723 -732, 2008.
- [18] S. M. Platt and N. I. Badler, "Animating facial expression," *Computer Graphics*, pp. 245-252, Aug. 1981.
- [19] K. Waters, "A muscle model for animating three-dimensional facial expression," in *Proc. the 14th Annual Conference on Computer Graphic and Interactive Techniques*, pp. 17-24, 1987.
- [20] D. Terzopoulos and K. Waters, "Physically-based facial modelling, analysis, and animation," *Journal of Visualization and Computer Animation*, vol. 1, no. 4, pp. 73-80, March, 1990.
- [21] Y. Lee, D. Terzopoulos, and K. Waters, "Constructing physics-based facial models of individuals," in *Proc. Graphic Interface 1993*, pp. 1-8, 1993.
- [22] W. Wang, W. Yan, Y. Xie, J. Qin, W.M. Pang, and P. A. Heng, "A physically-based modeling and simulation framework for facial animation," *2009 Fifth International Conference on Image and Graphics*, pp. 521-526, 2009.
- [23] S. Tang, H. Yan, and A. W. Liew, "A NURBS-based vector muscle model for generating human facial expressions," in *Proc. 2003 Joint Conference of the Fourth International Conference on Information, Communications and Signal Processing*, pp. 758-762, 2003.
- [24] S. Tang, A. W. Liew, and H. Yan, "Lip-Sync in human face animation based on video analysis and spline models," in *Proc. the 10th International Multimedia Modelling Conference*, pp. 102-108, 2004.
- [25] J. Noh and U. Neumann, "A survey of facial modeling and animation techniques," *Technical Report 99-705*, University of Southern California, 1998.

- [26] G. Borshukov, D. Piponi, O. Larsen, J. P. Lewis, and C. T. Lietz, "Universal capture - image-based facial animation for 'the matrix reloaded'," *ACM Siggraph 2005 Courses*, July 31-Aug. 04, 2005.
- [27] J. P. Lewis, J. Mooser, Z. Deng, and U. Neumann, "Reducing blendshape interference by selected motion attenuation," in *Proc. Siggraph 2005, Symposium on Interactive 3D Graphics and Games*, pp. 25-29, 2005.
- [28] Z. Deng, P. Y. Chiang, P. Fox, and U. Neumann, "Animating blendshape faces by cross-mapping motion capture data," *ACM SIGGRAPH Symposium on Interactive 3D Graphics and Games*, pp. 43-48, 2006.
- [29] F. I. Parke. "A model for human faces that allows speech synchronized animation," *Journal of Computers and Graphics*, vol. 1, no. 1, pp. 1-4, 1975.
- [30] J. P. Lewis and F. I. Parke, "Automated lip-synch and speech synthesis for character animation," in *Proc. CHI on Human Factors in Computing Systems and Graphic Interface*, pp. 143-147, 1987.
- [31] Y. Cao, P. Faloutsos, and F. Pighin, "Expressive speech-driven facial animation," *ACM Trans.on Graph.*, vol. 24, pp. 1283-1302., Oct. 2005.
- [32] C. Bregler, M. Covell, and M. Slaney, "Video rewrite: driving visual speech with audio," in *Proc. Siggraph 1997*, pp. 353-360, 1997.
- [33] Y. M. Chen, F. C. Huang, S. H. Guan, and B. Y. Chen, "Animating lip-synch characters with dominated animeme models," *IEEE Trans. on Circuits and Systems for Video Technology*, pp. 1344-1253, 2012.
- [34] L. Williams, "Performance-driven facial animation," in *Proc. 17th Computer Graphics and Interactive Techniques*, pp. 235-242, 1990.
- [35] B. Guenter, C. Grimm, D. Wood, H. Malvar, and F. Pighin, "Making faces," in *Proc. Siggraph 1998*, pp. 55-66, 1998.
- [36] C. Kouadio, P. Poulin, and P. Lachapelle, "Real-time facial animation based upon a bank of 3D facial expressions," in *Proc. Computer Animation*, pp. 128-136, 1998.
- [37] T. Weise, S. Bouaziz, H. Li, and M. Pauly, "Realtime performance-based facial animation," *ACM Transaction on Graphic*, vol. 30, issue 4, July 2011.



Heng Yu Ping is currently Ph.D. candidate in Multimedia System at Universiti Putra Malaysia (UPM). She received her Master in Computer Science from UPM in 2010 and Bac. in Computer Science (Multimedia) from UPM in 2008. Her research interests include facial modelling, facial animation, emotion recognition and physiological signal processing.



Lili Nurliyana Abdullah is currently an Assistant Professor in the Department of Multimedia, Faculty of Computer Science and Information Technology at Universiti Putra Malaysia. She received her Ph.D. degree in Information Science from Universiti Kebangsaan Malaysia, in 2007, M.S. degree in Engineering (Telematics) from University of Sheffield, United Kingdom, in 1996, and Bac. in Computer Science from UPM in 1990. Her research interests include multimedia system, video processing and retrieval, computer modelling and animation, image processing and computer games.



Puteri Suhaiza Sulaiman is currently a senior lecturer in the Department of Multimedia, Faculty of Computer Science and Information Technology at Universiti Putra Malaysia. She received her Ph.D. degree in Computer Graphic from Universiti Putra Malaysia, in 2010, M.S. in Computer Science from Universiti Teknologi Malaysia (UTM), in 2003, and Bac. in Computer Science from UTM, in 2000. Her research interests include computer graphics, geographical information systems, scientific visualization, medical visualization, flight simulation, virtual reality and computer games.



Alfian Abdul Halin is currently a senior lecturer in the Department of Multimedia, Faculty of Computer Science and Information Technology at Universiti Putra Malaysia. He received his Ph.D. degree in Computer Science from Universiti Sains Malaysia, in 2011, Master degree in Multimedia Computing from Monash University, Australia, in 2004, and B.S. in Information Technology from University Technology MARA, Malaysia, in 2000. His research interests include computer vision, image and audio processing, semantic concept detection, indexing and retrieval (in images and videos) and applied machine learning.