PATENT SPECIFICATION
Confidential

TITLE OF THE INVENTION

Methods and Apparatus for AI (Artificial Intelligence) Movie Producer System

INVENTORS

Shen Shengmei, Cai Mao, Tu Xiaoguang, Grace Ong
Pensees Pte Ltd
3 Fusionopolis Way, #14-23/24 Symbiosis, Singapore 138633

FIELD OF THE ART

The present invention relates to the field of movie or video creation with AI technologies. More specifically, the present invention relates to methods and apparatus for AI natural movie or natural video production by a AI system automatically with realistic human characters and faces, realistic animals, objects, scenery, realistic moving and actions, and realistic sound, and voice and speech conversation, based on given story in text, given novel book, given comic book, set of drawings, or given cartoon movie, or partial of them without complete content, furthermore an interactive user interface can be provided to allow human to involve in the production loop to create personalized movie or video content, to have the opportunity to modify the story content given, where the AI system also allows to be much creative to make a natural and complete movie or video with different starting and ending as well as the content in between even based on very limited storyline or incomplete text content or incomplete cartoon movie.

BACKGROUND ART

Nowadays in movie industry it is still very costly to make a movie, and the same for making a video, where you need to hire script writer, director, actor/actress, artist, musician, etc. To make a natural movie or video it also takes long time to complete, from several months to a couple of years.

Making a cartoon movie may cost less but the human character and scene are not natural, but more on animation style, less reality compared to a natural movie.

Recently AI technology can convert one image style to another, can create random content, can swap face, can render a given face to a video, can create action for a given person.
However it is still impossible to create a natural movie automatically without writing a scrip, engaging human and real actors/actress, designing and building various scenes, etc.

While on the other hand, there are a lot of story books available, in text form, or manga style like drawing pictures. Some are interested in reading the books or manga drawing pictures but it would be more demanding if such story books can be converted into natural movie or video to reflect the story content with AI technology.

SUMMARY OF THE INVENTION

Problems to be solved

The prior art as described above has the following problems.

Writing a script for a movie will take time while the available story books cannot be converted into movie content easily and automatically. On the other hand, even a movie script can be created from an available story books it still takes much of the time with a lot of resources including manpower resources and financial resources to make a natural movie from a movie script.

Making a carton movie is cheaper relatively but the human characters, objects and scene are not natural, less effective than a natural movie.

The current AI technology cannot be used to create natural movie content directly and automatically from a story book, a movie scrip, a manga, or a set of drawings although the current AI technology can change image from one style to another, can swap faces, can render a face to a video, can create certain type of images.

Means of Solving the Problems

The present invention defines a set of methods and apparatus to generate a natural movie or video content based on a given story book, a script or manga, or a set of drawings, which include Natural Language Processing-NLP to understand the story content written in natural language, followed by scene and background creation process with natural objects inside based on the NLP processing results, where the scene and background library, or the scene and background creation engine are provided; followed by natural human characters creation process like actors and actresses with natural faces and natural face expression as well as natural human attributes based on the NLP processing results, where the natural human character library, or the natural human character creation engine are provided; followed by natural animal creation process with natural attributes based on the NLP processing results, where the natural animal library, or the natural animal creation engine is provided; followed by natural facial expression changes and natural human movements, and natural object movements from one frame to another based on the NLP processing results, where the natural motion and video library, or the natural motion and video creation engine are provide. The technologies like Deep Learning Network, GAN (Generative Adversarial Network), Deep Reinforcement learning, or other technologies are used in the above natural movie creation process.

The present invention also defines one of the ways with AI technology to generate a video content or sequence with a given face image which could be existing or not

existing face in the world, where the face poses and expressions in the generated video follow the same poses and expressions in a given targeted video accordingly;

The present invention also defines one of the ways with AI technology to generate a video content or sequence with a given face image which could be existing or not existing face in the world, where the face poses and expressions in the generated video are predicted, or predicted with certain guideline, or randomly generated, or interpreted based on a given story text;

The present invention also defines one of the ways with AI technology to generate a video content or sequence with a given face image which could be existing or not existing face in the world, where the face poses and expressions in the generated video are created towards to a given certain pose and expression in between the generated images, or a given final pose and expression;

Effects of Invention

The present invention provides a significant improvement for the current movie industry, to create natural movie or natural video in cheap and fast way by utilizing AI technology.
By giving a story book, or movie script, or a set of drawings, or a cartoon movie a natural movie or video can be produced automatically with AI technology.

By doing this way a lot of story books and story scripts can be converted into a natural movie or video quickly and effectively for people to enjoy in multiple ways not only in text form but also in movie or video style with vivid visual pictures and various sound effects;
By doing this way more people can involve in movie creation and production;
By doing this way more personalized movie and video can be created;
By doing this way story book writer can be facilitated to make their own natural movie or video besides writing their story books;
By doing this way more and more movie directors and creators to make natural movie would rise up;
By doing this way natural video content can be created in a cheap and quick way based on a simple script or story line;

With the current invention with AI technology a video content or sequence can be created using desired faces with desired poses and expressions which can be driven by text content, voice content, visual content, or just simple guidline.

BEST MODE FOR CARRYING OUT THE INVENTION

With reference to the accompanying drawings, the preferred embodiment of the present invention are described as following.

FIG. 1 shows the general diagram for this invented AI Movie Producer system, which is described here.

The AI system receives a story book, or movie script, or manga book, or a set of drawings, or a cartoon movie, or an animation in the module 1.1.
The AI system creates a natural movie or video automatically and effectively with or without human involvement or simple interaction, based on the given story. The said created natural movie or a natural video can be with different movie and video styles, different starting, different endings, different natural human characters, different natural scenes, different natural background music, different natural voices and conversations, etc., depending on human movie producer's desire and intention with the help of interface between human producer and the invented AI Movie Producer system.

FIG. 2 shows the detailed diagram of this invented AI Movie Producer system based on a story in text, manga, a set of drawings, which is described here.

The AI system receives the input story content in text form or manga form in the module of 2.1.
The AI system in the module of 2.2 reads through the story and digests the story line and contents in great details with AI and NLP (Natural Language Processing) technology, to understand the scene, identify the human characters and other objects with their attributes, their actions and movements in the content of the story.

The AI system in the module of 2.3 creates natural scenes and backgrounds with music and sounds corresponding to the content of the story with natural seas, mountains, rivers, trees, streets, roads, fields, parks and other objects etc., with their attributes like color, size, shape, type, pattern, etc. The AI system in the module of 2.8 can make change to the attributes to the created scene and backgrounds, the created objects, and the positioning and layout among of them.

The AI system in the module of 2.4 creates natural human actors and actresses corresponding to the content of the story with natural face and natural human attributes. The AI system in the module of 2.9 can make change to the attributes of the created actors and actresses including but not limited to their ages, face skins, face shapes, hair styles, wearing accessories like glasses, masks, scarfs, hats, body size and shapes, cloths with different colors and patterns, and other human attributes like carrying bags, holding baby, riding on the motorbike, etc.

The AI system in the module of 2.5 creates natural animals with natural sounds corresponding to the content of the story with their attributes like color, pattern of the skin, size and shape, etc.

The AI system in the module of 2.6 creates and changes natural human facial expression, natural action and movement, and natural voices and sounds corresponding to the content of the story with natural and smooth transition from ne frame to another.

The AI system in the module of 2.7 creates and changes the object movements and sounds corresponding to the content of the story with natural and smooth transition from one frame to another.

FIG. 3 shows the detailed diagram of this invented AI Movie Producer system based on a story in a cartoon movie, or an animation, which is described here.

The AI system receives the input story content in a cartoon movie, or an animation in the module of 3.1.
The AI system in the module of 3.2 reads through the story and digests the story line and contents in great details with AI image and video understanding and analytics technology, to understand the scene, identify the human characters and other objects with their attributes, their actions and movements in the content of the story.

The AI system in the module of 3.3 creates natural scenes and backgrounds with music and sounds corresponding to the content of the story with natural seas, mountains, rivers, trees, streets, roads, fields, parks and other objects etc., with their attributes like color, size, shape, type, pattern, etc. The AI system in the module of 3.8 can make change to the attributes to the created scene and backgrounds, the created objects, and the positioning and layout among of them.

The AI system in the module of 3.4 creates natural human actors and actresses corresponding to the content of the story with natural face and natural human attributes. The AI system in the module of 3.9 can make change to the attributes of the created actors and actresses including but not limited to their ages, face skins, face shapes, hair styles, wearing accessories like glasses, masks, scarfs, hats, body size and shapes, cloths with different colors and patterns, and other human attributes like carrying bags, holding baby, riding on the motorbike, etc.

The AI system in the module of 3.5 creates natural animals with natural sounds corresponding to the content of the story with their attributes like color, pattern of the skin, size and shape, etc.

The AI system in the module of 3.6 creates and changes natural human facial expression, natural action and movement, and natural voices and sounds corresponding to the content of the story with natural and smooth transition from ne frame to another.

The AI system in the module of 3.7 creates and changes the object movements and sounds corresponding to the content of the story with natural and smooth transition from one frame to another.

FIG. 4 shows the diagram for the invented AI Movie Producer system with the user interface, which is described here.

The AI system receives a story in text, manga, a set of drawing, a cartoon movie, or an animation in the module of 4.1.

Human as a user with their computer, laptop, pad, or mobile phone to interact with AI Movie Producer system which provides a User Interface, in the module of 4.2.

The said User Interface in the module of 4.3 allows the user to select, search their desired items from the given Library in the module of 4.4, or even deep involvement in the movie creation process on the top of automatic function, including but not limited to penning, drawing, writing, editing, recording, inserting to make their desired and personalized movie. The desired items include but not limited to scenes with various scene structures, with various objects, with various styles; human characters, with various human faces, with various facial expression, with various human attributes, with various actions, motions and voices; objects with various attributes and motions.

The AI system provides the AI Movie Producer library besides the fully automatic movie creation function, in the module of 4.4, where huge amount of the items provided to be used for movie creation are pre-created and on-line learnt and updated by deep learning network, GAN (Generative Adversarial Network), deep reinforcement learning, other technologies, and such items includes but not limited to scenes with various scene structures, various objects, various styles; human characters, with various human faces, various facial expression, various human attributes, various actions, motions and voices; objects with various attributes and motions.

With reference to the accompanying drawings of Fig 5 and 6, the preferred embodiments of the present invention are described as following.

A novel framework for high-realistic and identity-preserving face video content generation from a single face image is invented, where the facial dynamics are taken as conditions to guide face generation by a prior-guided face generator. A 3D dynamic generation network is invented to predict spatiotemporal consistent 3D dynamic sequence. A 3D down-sampled texture mapping is designed to render the 3D face dynamics as prior knowledge, which covers rich content information of the source image to guide face generation. Face video generation in both stochastic and controllable manner is covered, where the former predicts a reasonable future face frames, while the later can be controlled by rules either defined in a given targeted video with designed face frames or controlled by certain guideline so to generate face frames towards a given target.

Fig 5 shows the diagram for the invented video generation method based on a given image with 3D reconstruction and 3D Vertex.

Given a face image as source face, the 3D face shape in the module of 5.2 and 3D vertices in the module of 5.3 can be obtained by estimating the coefficients of a 3D Morphable Model (3DMM) in the module of 5.1. By changing the 3DMM coefficients as 3Dc1,3Dc2,...,3Dck in the module of 5.4, the 3D face can be re-shaped and its surface can be rendered based on 3D down-sampled texture mapping in the module of 5.5. The sampled textures (P0,...,Pk) are used for guiding face generation.

3D face fitting is used here to estimate the 3D face shape, expression and camera projection matrix from a given input face image. We estimate the 3D face from a single face image based on 3D Morphable Model (3DMM), in which the 3D vertices of a 2D face image can be expressed as linear combination over a set of PCA (Principle Component Analysis) bases.

FIG. 6 shows the diagram for the invented video generation method with LSTM for prediction and encoder and decoder for generation network

3D Dynamic Prediction(3DDP) Network (a) is shown in the module of 6.1.
The future 3D dynamic (3Ddi) sequence of a source face is predicted by the 3DDP network. The source face along with the generated 3Ddi are used to generate the 3D prior(Pi), which are used to guide face generation in the Face generation network (b) in the module of 6.2.

The source face image Is in the module of 6.1 is used to render the predicted 3D dynamics based on 3D down-sampled texture mapping. Such 3D down-sampled textures are the 3D priors that we used for guiding face generation. The face generation model is a conditional GAN (cGAN), which contains a prior-guided face generator G for high-realistic face generation conditioned on the predicted 3D prior sequence, and a dual-input discriminator D to distinguish the real/fake of the outputs regarding the target 3D priors. Fig. 3 (b) shows the architecture of our generation model.

Face generation network in the module of 6.2 takes in three inputs, including a source face image Is, Is's corresponding 3D prior Ps and a target 3D prior Pt. Instead of using Pt to directly guide face generation, we feed Ps and Pt into the encoder and do subtraction to obtain the prior residual. Therefore, the target face image can be predicted from Is conditioned on the prior residual
The prior-guided generator contains an encoder E and a decoder F. F takes the concatenated representation of the input face image and the prior residual between the source and target priors as the input. D takes the target prior as condition and pairs the real/fake images as dual input.

Industrial Applicability

The present invention is a new and effective way for entertainment industry to make new movie and video content, an efficient and cost saving way for the existing movie industry to make a new movie, a quick and cheap way for video content creation industry, where natural movie and video content can be created directly and automatically from a story book, a movie scrip, a set of drawing with AI technology and system to reduce huge cost and time, to provide new user experience to reflect user's desire and interest.

With AI Movie Producer system, a natural movie characters can be created, a natural movie scene or background can be created, a natural facial expression, natural human motion can be produced, natural sounds, voices, and conversations can be produced.

With the invented video content generation with AI technology, a desired video with the desired realistic human face and the desired face pose, face shape and face expression can be generated, which could be part of a AI generated movie.

BRIEF DESCRIPTION OF THE DRAWINGS

FIG. 1 shows a general diagram for the invented AI Movie Producer system to create a natural movie in an efficient way.
FIG. 2 shows the detail diagram for the invented AI Movie Producer system based on a story in text, manga, a set of drawings.
FIG. 3 shows the detail diagram for the invented AI Movie Producer system based on a story in a cartoon movie, or an animation.
FIG. 4 shows the diagram for the invented AI Movie Producer system with the user interface.
FIG. 5 shows the diagram for the invented video generation method based on a given image with 3D reconstruction and 3D Vertex:
FIG. 6 shows the diagram for the invented video generation method with LSTM for prediction and encoder and decoder for generation network:

WHAT IS CLAIMED IS:

1.Methods and apparatus for AI Movie Producer system, comprising:

Receiving a story line in text, manga, or in set of drawings,

Reading through the story line and digesting the story content,

Creating a natural movie or video automatically based on the content of the said story in text or in set of drawings.

2. Methods and apparatus for AI Movie Producer system, in the claim 1 where Reading through the story line and digesting the story content, further comprising:

Understanding the scene with the corresponding objects and attributes based on the content of the said story in text or in set of drawings,

Identifying the human characters with the corresponding attributes based on the content of the said story in text or in set of drawings,

Identifying the animals with the corresponding attributes based on the content of the said story in text or in set of drawings,

Identifying the facial expression, human actions and movements based on the content of the said story in text or in set of drawings,

Identifying the other objects (car for example) with the corresponding attributes and motion based on the content of the said story in text or in set of drawings.

3. Methods and apparatus for AI Movie Producer system, in the claim 1 where Creating a natural movie or video automatically based on the content of the said story in text or in set of drawings, further comprising:

Creating the natural scenes with natural objects like seas or mountains or tress or buildings etc. with their attributes, and changing their positions and layout among of them, based on the content of the said story in text or in set of drawings,

Creating the natural actors and actresses with natural face expressions, natural human attributes, natural human actions, and natural voices and conversations, and making the change to the said natural face expressions, natural human attributes, natural human actions, and natural voices and conversations with smooth and natural transition in between frames, based on the content of the said story in text or in set of drawings,

Creating the natural animals with their attributes and actions, and making the change to the said attributes and actions with smooth and natural transition in between frames, based on the content of the said story in text or in set of drawings,

Creating the other natural objects (for example cars) with their attributes and actions, and making the change to the said attributes and actions with smooth and natural transition in between frames, based on the content of the said story in text or in set of drawings.

4. Methods and apparatus for AI Movie Producer system, in the claim 1 where Creating a natural movie or video automatically based on the content of the said story in text or in set of drawings, further comprising:

Creating the natural scenes with natural objects like seas or mountains or tress or buildings etc. with their attributes, and changing their positions and layout among of them, based on but not exactly the same as the content of the said story in text or in set of drawings,

Creating the natural actors and actresses with natural face expressions, natural human attributes, natural human actions, and natural voices and conversations, and making the change to the said natural face expressions, natural human attributes, natural human actions, and natural voices and conversations with smooth and natural transition in between frames, based on but not exactly the same as the content of the said story in text or in set of drawings,

Creating the natural animals with their attributes and actions, and making the change to the said attributes and actions with smooth and natural transition in between frames, based on but not exactly the same as the content of the said story in text or in set of drawings,

Creating the other natural objects (for example cars) with their attributes and actions, and making the change to the said attributes and actions with smooth and natural transition in between frames, based on but not exactly the same as the content of the said story in text or in set of drawings.

5. Methods and apparatus for AI Movie Producer system, in the claim 1 where Creating a natural movie or video automatically based on the content of the said story in text or in set of drawings, further comprising:

Having a function to interact with users with AI Movie Producer system for their selections from the provided library which is the part of AI Movie Producer system, and such selections including but not limited to scene structure and their attributes, human characters, human faces, human attributes, animals and their attributes, objects and their attributes, human, animal and object actions and movements, part of video with certain motions,

Having a function to interact with users with AI Movie Producer system for their search for their desired items with key words from the provided library which is the part of AI Movie Producer system, and such items including but not limited to scene structure and their attributes, human characters, human faces, human attributes, animals and their attributes, objects and their attributes, human, animal and object actions and movements, part of video with certain motions,

Having a function to interact with users with AI Movie Producer system for their further involvement including but not limited to drawing, penning, writing, editing with computer or mobile phone user interfaces, sound and voice recording, then to insert into the scene and image and sound track, or modify the scene and image or sound track based on their interest and favor, and such interests including but not limited to scene structure and their attributes, human characters, human faces, human attributes, animals and their attributes, objects and their attributes, human, animal and object actions and movements, part of video with certain motions.

6. Methods and apparatus for AI Movie Producer system, comprising:

Receiving a story line in cartoon movie, or animation,

Reading through the story line and digesting the story content,

Creating a natural movie or video automatically based on the content of the said story in the said cartoon movie, or animation.

7. Methods and apparatus for AI Movie Producer system, in the claim 6 where Reading through the story line and digesting the story content, further comprising:

Understanding the scene with the corresponding objects and attributes based on the content of the said story in cartoon movie, or animation,

Identifying the human characters with the corresponding attributes based on the content of the said story in cartoon movie, or animation,

Identifying the animals with the corresponding attributes based on the content of the said story in cartoon movie, or animation,

Identifying the facial expression, human actions and movements based on the content of the said story in cartoon movie, or animation,

Identifying the other objects (car for example) with the corresponding attributes and motion based on the content of the said story in cartoon movie, or animation.

8. Methods and apparatus for AI Movie Producer system, in the claim 6 where Creating a natural movie or video automatically based on the content of the said story in cartoon movie, or animation, further comprising:

Creating the natural scenes with natural objects like seas or mountains or tress or buildings etc. with their attributes, and changing their positions and layout among of them, based on the content of the said story in cartoon movie, or animation,

Creating the natural actors and actresses with natural face expressions, natural human attributes, natural human actions, and natural voices and conversations, and making the change to the said natural face expressions, natural human attributes, natural human actions, and natural voices and conversations with smooth and natural

transition in between frames, based on the content of the said story in cartoon movie, or animation,

Creating the natural animals with their attributes and actions, and making the change to the said attributes and actions with smooth and natural transition in between frames, based on the content of the said story in cartoon movie, or animation,

Creating the other natural objects (for example cars) with their attributes and actions, and making the change to the said attributes and actions with smooth and natural transition in between frames, based on the content of the said story in cartoon movie, or animation.

9. Methods and apparatus for AI Movie Producer system, in the claim 6 where Creating a natural movie or video automatically based on the content of the said story in cartoon movie, or animation, further comprising:

Creating the natural scenes with natural objects like seas or mountains or tress or buildings etc. with their attributes, and changing their positions and layout among of them, based on but not exactly the same as the content of the said story in cartoon movie, or animation,

Creating the natural actors and actresses with natural face expressions, natural human attributes, natural human actions, and natural voices and conversations, and making the change to the said natural face expressions, natural human attributes, natural human actions, and natural voices and conversations with smooth and natural transition in between frames, based on but not exactly the same as the content of the said story in cartoon movie, or animation,

Creating the natural animals with their attributes and actions, and making the change to the said attributes and actions with smooth and natural transition in between frames, based on but not exactly the same as the content of the said story in cartoon movie, or animation,

Creating the other natural objects (for example cars) with their attributes and actions, and making the change to the said attributes and actions with smooth and natural transition in between frames, based on but not exactly the same as the content of the said story in cartoon movie, or animation.

10. Methods and apparatus for AI Movie Producer system, in the claim 6 where Creating a natural movie or video automatically based on the content of the said story in cartoon movie, or animation, further comprising:

Having a function to interact with users with AI Movie Producer system for their selections from the provided library which is the part of AI Movie Producer system, and such selections including but not limited to scene structure and their attributes, human characters, human faces, human attributes, animals and their attributes, objects and their attributes, human, animal and object actions and movements, part of video with certain motions,

Having a function to interact with users with AI Movie Producer system for their search for their desired items with key words from the provided library which is the part of AI Movie Producer system, and such items including but not limited to scene structure and their attributes, human characters, human faces, human attributes, animals and their attributes, objects and their attributes, human, animal and object actions and movements, part of video with certain motions,

Having a function to interact with users with AI Movie Producer system for their further involvement including but not limited to drawing, penning, writing, editing with computer or mobile phone user interfaces, sound and voice recording, then to insert into the scene and image and sound track, or modify the scene and image or sound track based on their interest and favor, and such interests including but not limited to scene structure and their attributes, human characters, human faces, human attributes, animals and their attributes, objects and their attributes, human, animal and object actions and movements, part of video with certain motions.

11.Methods and apparatus for AI Movie Producer system, comprising:

Inputting an available single image or a set of face images, or creating realistic human face image with AI technology as the source image,

Creating 3D Morphable Model (3DMM) coefficients on the above source image,

Estimating 3D face shape and 3D vertices from the said 3DMM coefficients,

Re-shaping 3D shapes by changing 3DMM coefficients,

Rendering 3D shapes with textures for guiding face generation in video content creation process.

12. Methods and apparatus for AI Movie Producer system, in the claim 11 where Re-shaping 3D shapes by changing 3DMM coefficients, further comprising:

Predicting 3DMM coefficients from LSTM network (Long Short-Term Memory Network) with designed condition or random, or from targeted face frames in a video sequence.