CLAIMS

1. A method for video communications implemented by a computer system, the method comprising:

implementing a 3D virtual environment configured to be accessed by a plurality of client devices, each having a corresponding user graphical representation within the 3D virtual environment, each user graphical representation having a corresponding virtual camera comprising a view of the 3D virtual environment presented through a display to the corresponding users of the client devices, wherein the computer system enables exchanging video and audio data between users of the 3D virtual environment to provide video communications in the 3D virtual environment;

receiving head tracking metadata of a first client device that comprises 6 degrees of freedom head tracking information generated in response to tracking movement of key facial landmarks obtained from video of a face of a first user of the first client device, wherein a first virtual camera comprises a view of the 3D virtual environment presented to the first user;

associating the view of the first virtual camera to the coordinates of the key facial landmarks;

tracking movement of the key facial landmarks in the 6 degrees of freedom based on the movement of the head of the first user;

adjusting the position and orientation of the first virtual camera based on the tracked movement of the key facial landmarks;

identifying graphical elements within a field of view of a second virtual camera comprising a view of the 3D virtual environment presented to a second user of a second client device, wherein the identified graphical elements comprise a user graphical representation of the first user;

sending, to the second client device, the head tracking metadata of the first client device and the identified graphical elements;

using, by the second client device, the head tracking metadata to adjust movement of a head of the user graphical representation of the first user; and

rendering and presenting a modified presentation of the view of the second virtual camera including the user graphical representation of the first user and corresponding graphical elements.

2. The method of claim 1, wherein the head tracking metadata is sent to the second client device via a network using a video channel.

3. The method of claim 1, wherein threshold values for one or more of the degrees of freedom limit the adjustments in the position and orientation of the virtual camera.

4. The method of claim 3, wherein the movement of the head in each of the degrees of freedom results in a corresponding change of the virtual camera only after reaching the threshold value.

5. The method of claim 3, wherein the movement of the virtual camera associated with the movement of the head in each of the degrees of freedom is stopped after reaching the threshold value.

6. The method of claim 1, comprising applying a multiplier value to one or more of the degrees of freedom, resulting in an increased movement of the virtual camera with respect to each movement of the head.

7. The method of claim 1, comprising inverting the direction of the adjustments of the position and orientation of the virtual camera with respect to the tracked movements of the key facial landmarks.

8. The method of claim 1, wherein the 3D virtual environment includes positions for the user graphical representations and their corresponding virtual cameras arranged in a geometry, wherein the adjustments of the position and orientation of the virtual cameras are controlled based on a horizontal rotation or lateral movement of the key facial landmarks so that the virtual cameras are moved on a predetermined path arranged in the geometry.

9. The method of claim 1, comprising associating one or more axes of movement of the first virtual camera to one or more head movements.

10. A system for video communications, comprising:

at least one computer of a computer system with memory and at least one processor implementing a 3D virtual environment configured to be accessed by a plurality of client devices, each having a corresponding user graphical representation within the 3D virtual environment, each user graphical representation having a corresponding virtual camera comprising a view of the 3D virtual environment presented through a display to the corresponding users of the client devices;

a data exchange management module implemented in memory and configured to receive head tracking metadata of a first client device that comprises 6 degrees of freedom head tracking information generated in response to tracking movement of key facial landmarks obtained from video of a face of a first user of the first client device, wherein a first virtual camera comprises a view of the 3D virtual environment presented to the first user; and

a machine vision module implemented in memory and configured to identify graphical elements within a field of view of a second virtual camera comprising a view of the 3D virtual environment presented to a second user of at least one second client device, wherein the identified graphical elements comprise a user graphical representation of the first user, wherein the machine vision module is further configured to send to the second client device, via the data exchange management module, the head tracking metadata of the first client device and the identified graphical elements, and wherein the system enables exchanging video and audio data between users of the 3D virtual environment to provide video communications in the 3D virtual environment;

wherein the machine vision module is further configured to associate the view of the first virtual camera to the coordinates of the key facial landmarks, track movement of the key facial landmarks in 6 degrees of freedom based on the movement of the head of the first user, and adjust the position and orientation of the first virtual camera based on the tracked movement of the key facial landmarks;

wherein the second client device is configured to:

receive the head tracking metadata to adjust movement of a head of the user graphical representation of the first user; and

render and present a modified presentation of the view of the second virtual camera including the user graphical representation of the first user and corresponding graphical elements.

11. The system of claim 10, wherein the head tracing metadata is sent to the second client device via a network using a video channel.

12. The system of claim 10, wherein the key facial landmarks along with their coordinates are obtained and identified from color image frames of the first user.

13. The system of claim 10, wherein the machine vision module implements threshold values for one or more of the degrees of freedom to limit the adjustments in the position and orientation of the virtual camera.

14. The system of claim 13, wherein the machine vision module implements the threshold values such that the movement of the head in each of the degrees of freedom results in a corresponding change of the virtual camera only after reaching the threshold value.

15. The system of claim 13, wherein the machine vision module implements the threshold values such that the movement of the virtual camera associated with the movement of the head in each of the degrees of freedom is stopped after reaching the threshold value.

16. The system of claim 10, wherein the machine vision module implements a multiplier value for one or more of the degrees of freedom, resulting in an increased movement of the virtual camera with respect to each movement of the head.

17. The system of claim 10, wherein the 3D virtual environment includes positions for the user graphical representations and their corresponding virtual cameras arranged in a geometry, and wherein the adjustments of the position and orientation of the virtual cameras are controlled based on a horizontal rotation of the key facial landmarks so that the virtual cameras are moved on a predetermined path arranged in the geometry.

18. A non-transitory computer-readable medium having stored thereon instructions configured to cause at least one server computer comprising a processor and memory to perform steps comprising:

implementing a 3D virtual environment configured to be accessed by a plurality of client devices, each having a corresponding user graphical representation within the 3D virtual environment, each user graphical representation having a corresponding virtual camera comprising a view of the 3D implementing a 3D virtual environment configured to be accessed by a plurality of client devices, each having a corresponding user graphical representation within the 3D virtual environment, each user graphical representation having a corresponding virtual camera comprising a view of the 3D virtual environment presented through a display to the corresponding users of the client devices, wherein the at least one server computer enables exchanging video and audio data between users of the 3D virtual environment to provide video communications in the 3D virtual environment;

receiving head tracking metadata of a first client device that comprises head tracking information generated in response to tracking movement of key facial landmarks obtained from video of a face of a first user of the first client device;

associating the view of the first virtual camera to the coordinates of the key facial landmarks;

tracking movement of the key facial landmarks in 6 degrees of freedom based on the movement of the head of the first user;

adjusting the position and orientation of the first virtual camera based on the tracked movement of the key facial landmarks;

identifying graphical elements within a field of view of a virtual camera of a second client device, wherein the identified graphical elements comprise a user graphical representation of the first user;

sending, to the second client device, the head tracking metadata of the first client device and the identified graphical elements;

using, by the second client device, the head tracking metadata to adjust movement of a head of the user graphical representation of the first user; and

rendering and presenting a modified presentation of the view of the second virtual camera including the user graphical representation of the first user and corresponding graphical elements.

19. The non-transitory computer-readable medium of claim 18, wherein the head tracking metadata is sent to the second client device via a network using a video channel.