Title:
Generating 3D CAD Art from Human Gestures using Kinect Depth Sensor

Authors:
Miri Weiss Cohen, Miri@braude.ac.il, Braude College of Engineering, Israel
Alex Frid, Alexf@braude.ac.il, Braude College of Engineering, Israel
Mark Melin, drawkubes@gmail.com, Braude College of Engineering, Israel
Vladimir Ladijenski, Vovchech83@gmail.com, Braude College of Engineering, Israel

Keywords:
3D CAD art, kinect sensor, natural user interface (NUI), computer art

DOI: 10.14733/cadconfP.2014.161-163

Introduction:
Microsoft Kinect was originally designed as a motion-sensing device developed as an Xbox console controller that facilitates interaction using gestures and body motion, but its applicability goes beyond the gaming domain as it can be used in robotics, medicine and arts [1,5,6]. In this study we propose the generation of visual interactive art with 3D geometric features using Microsoft Kinect sensor. Natural human movement and gesture recognition are used for creation and interaction with various objects in 3D space for art design. The art design results could then be captured to yield a static art form, a dynamic video art design or an interactive reactive video art, all providing another dimension to the art domain.

Main idea:
An algorithm has been developed that uses the Kinect output as it records movement to create a structure that will accurately represent the movement captured and to preserve the sequence of movements either for future works or to apply transformations on existing structures. The last stage is to convert the structured data into its final state as art design. The proposed algorithm consists of three main stages (I) time dependent depth data acquisition and joints identification, followed by (II) weighted undirected graph generation that passed to (III) graph scanning algorithm with visual conversion and post processing:

- *Stage 1 - Depth sensor data acquisition:* At this stage Microsoft Kinect sensor is used, which is a motion capture device developed by Microsoft for the Xbox 360 video game console and Windows PCs. The device consists of a webcam, a depth camera and a microphone array that enables users to control and interact with the console by means of gestures and voice commands. As the body parts are identified [2,4], the sensor positions the joints with the highest matched data, thus enabling the movement of the defined joints to track the movement of the whole body (Fig. 1).
- *Stage 2 - Depth data to graph transformation:* Given are 10-20 joints of the human skeleton per frame (in time *t*) in in three dimensional space. . Hence each joint is represented by *(x,y,z,t)* where *x,y,z* are the location coordinates of the joint and *t* is the time. Between each joint (x_i, y_i, z_i) where *i* is a unique joint number there will be an edge (clique) at time *t=0* with a weight function w' which will be a random number between 1 and the number of joints. Between each joint (x_i, y_i, z_i) at time *t* and the same joint (x_i, y_i, z_i) at *t+1* will be a double weighted edge. The first weight function w_1 will be calculated as follows: $w_1 = \sqrt{\Delta x^2 + \Delta y^2}$.
  And the second weight function will be calculated as $w_2 = (x_i - x_{i+1}, y_i - y_{i+1})$.
  The output will be a graph of joints over time (Fig. 2).

(a) (b)

Fig. 1: Human body recognition: (a) the body is recognized; (b) body parts are recognized and associated with joints [4].

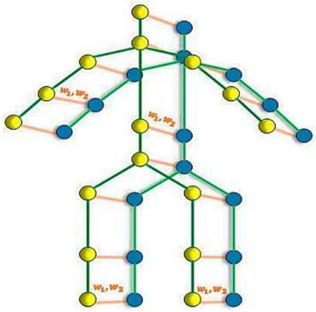
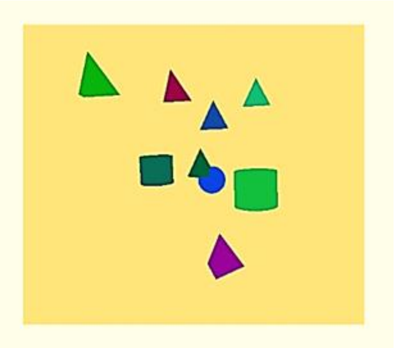


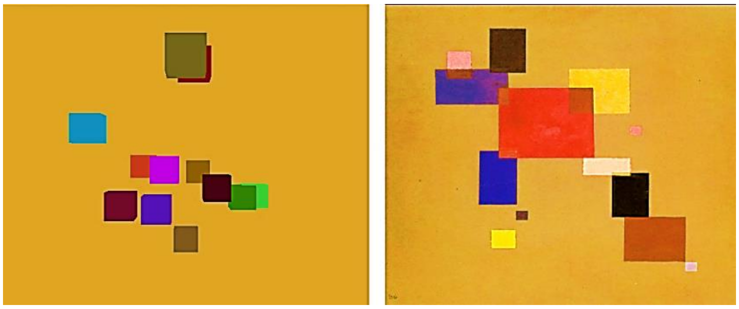
Fig. 2: Illustration of graph *G(V,E).*

- *Stage 3 - Graph to visual transformation:* At this stage, a graph scanning algorithm, namely, Depth Limited Search (DLS) is applied in order to choose the optimal path for the visual conversion. We choose a random number of 3D shapes for each shape and then choose a joint to start the search algorithm. The starting joint will also define the initial color of the shape as follows: The (x_i, y_i, z_i) coordinates of each joint correspond to R, G, and B, respectively. The weight w' sum on the edges is calculated, resulting in a corresponding predefined shape. After choosing a starting joint, the system calculates w_2. The movement of each shape is defined by the direction vector w_2 and defines the speed. Moreover, choice of a different joint will lead to calculation of w'. Data acquisition is a sequential process which lasts for a constant time t. During this time the second and third stages are applied constantly, and the result is immediately reflected in the shown design, straight after post processing step which adds perspective correction making the process interactive and dynamic. Figures 3 and 4 show two static snapshots of the system. Figure 3 shows a very basic set of 3D geometric forms in a specific t. Figure 4 is an attempt to transform *Thirteen-Rectangles-1930* [3], an existing (2D) abstract art piece by Wassily Kandinsky used as the inspiration for a 3D interactive video art.

Conclusions:

The proposed 3D art design system generates an abstract and unique geometrical design based on natural human gestures and movements. Shapes, colors, placements and speed are defined, captured and translated to art forms. Because each person is unique and each movement is different and cannot be replicated precisely, the process will yield a new unique design each time. Interestingly, though the design process is fully computerized, the generated structures are personal, natural and very similar to those of humans.





(a) (b)

Fig. 3: A snapshot of 3D art design.

Fig. 4: (a) Our system design snapshot, (b) Wassily-Kandinsky [3].

Proceedings of CAD'14, Hong Kong, June 23-26, 2014, 161-163

References:

[1] Castro, B.-P.; Velho, L.; Kosminsky, D.: Integrate: digital art using body interaction, Computational Aesthetics in Graphics, Visualization and Imaging, Cunningham D. and House D. (Editors), 2012.

[2] Jana, A.: Kinect for Windows SDK Programming Guide, Packet Publishing, Birmingham, UK, December, 2012. ISBN: 1849692386.

[3] Kandinsky, W.: Thirteen-Rectangles-1930, Musée des Beaux Arts, Nantes, France.

[4] Le, T.-L.; Nguyen, M.; Nguyen, T.: Human posture recognition using human skeleton provided by Kinect, 2013 International Conference on Computing, Management and Telecommunications (ComManTel), 2013, 340,345. http://dx.doi.org/10.1109/ComManTel.2013.6482417

[5] Sanna, A.; Lamberti, F.; Paravati, G.; Rocha, F. –D.: A kinect-based interface to animate virtual characters, J. of Multimodal User Interfaces, 7, 2013, 269–279. http://dx.doi.org/10.1007/s12193-012-0113-9

[6] Rodrigues, D.-G; Grenader, G.: MotionDraw: A Tool for Enhancing Art and Performance Using Kinect, in CHI '13 Extended Abstracts on Human Factors in Computing Systems, NY, USA, 2013, 1197–1202.