# Head Pose 3D Data Web-based Visualization

Grigorios Kalliatakis*
Department of Informatics Engineering,
Technological Educational Institute of Crete, Greece

Georgios Triantafyllidis†
Mediology Section,
AD: MT Aalborg University Copenhagen, Denmark

Nikolaos Vidakis‡
Department of Informatics Engineering,
Technological Educational Institute of Crete, Greece

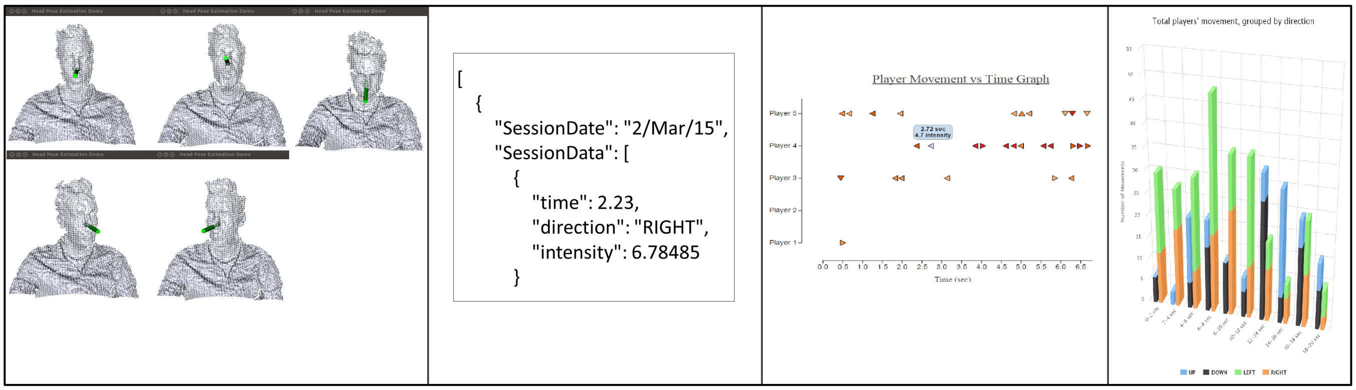


**Figure 1 :** The system overview (a) Head Pose Estimation results (b) JSON structure for storing the results (c) 2D Scatterplot Visualization for players' movement vs time (d) 3D Bar Chart Visualization for players' movement grouped by direction

## Abstract

An approach build on discriminative random regression forests was followed in order to achieve fast, accurate and reliable estimation of head pose in uncontrolled environment. Data representing the changes of a person's head direction, concerning two Degrees of Freedom (DOF), pitch and yaw, are collected and stored adopting a lightweight data exchange format (JavaScript Object Notation-JSON). After that, a web visualization approach is proposed in order to improve the understanding and the analysis of the captured 3D data.

**Keywords:** Head Pose Estimation, Web-based Data Visualization, Data Compilation, Kinect

## 1 Introduction

Because of its numerous applications, head pose estimation has drawn great attention from academia and a variety of techniques have been reported in the literature [Chutorian 2009].

*e-mail: gkalliatakis@yahoo.gr
†e-mail: gt@create.aau.dk
‡e-mail: nv@ie.teicrete.gr


Web3D '15, June 18 - 21, 2015, HERAKLION, Greece
ACM 978-1-4503-3647-5/15/06.
http://dx.doi.org/10.1145/2775292.2778304

With the recent technological advancements of depth sensors, it is possible to perform data collection in terms of head movement for subsequent analysis. Providing an objective assessment and evaluation of those findings can lead to valuable conclusions regarding the overall experience of user in many applications (e.g. in the case of educationally -oriented tasks in terms of evaluating user training and ludology experience in serious games).
Efficient visualization of such data can play a major part in that kind of assessment by creating encodings of data into visual channels that people can view and understand comfortably. The proposed scheme presented in this paper consists of three primary parts as shown in Figure 1. First the 3D head pose estimation events are obtained as shown in Figure 1(a). Then the data for every user are stored in a JSON file for offline usage, Figure 1(b), before creating the 2D and 3D visualizations with scatter plot and 3D columns, Figure 1(c) and 1(d) respectively. The principal objective of the proposed scheme is considered to be the acquisition of efficient and user-friendly visualizations in order to improve the understanding and the analysis of the captured data, which can be accessed easily through the web.

## 2 Head Pose Estimation Framework

The method proposed by [Fanelli et al 2011] is utilized in our work as it is regarded to be suitable for real time 3D head pose estimation, considering its robustness to the poor signal-to-noise ratio of current consumer depth cameras like Microsoft Kinect sensor. An extracted patch from a depth image is sent through all trees in the forest. The patch is evaluated at each node according to the stored binary test and passed either to the right or left child until a leaf node is reached [Fanelli et al. August 2012], at which point it is classified and only if this classification outcome is positive (head leaf), a Gaussian distribution is recaptured and then used for casting a vote in a multidimensional continuous space which is stored at the leaf. The system is able to perform on a frame-by-frame basis while it runs in real time without the need of initialization

The green cylinder encodes both the estimated head center and direction of the face as shown in Figure 1(a).

## 3 Data Compilation

Data representing the changes of a person's head direction, concerning two DOF, pitch and yaw, had to be obtained. The current streaming frame of Microsoft Kinect sensor had to be examined in contrast with the immediate previous (returned by the forest) for tracing the direction shift. A threshold was experimentally set around 4 for controlling the direction change as shown in (1) - (6) of a person's head in relation to the frames difference.

$$pitch\_difference = pitch_{t-1} - pitch_t \quad (1)$$
$$yaw\_difference = yaw_{t-1} - yaw_t \quad (2)$$

$$UP = pitch\_difference > threshold \quad (3)$$
$$DOWN = pitch\_difference < threshold \quad (4)$$

$$LEFT = yaw\_difference > threshold \quad (5)$$
$$RIGHT = yaw\_difference < threshold \quad (6)$$

Regarding the storage of the obtained data, JavaScript Object Notation (JSON) format was used mainly because of its lightweight nature, convenience in writing and reading and more importantly, its usefulness in generating and parsing tasks in various Ajax applications as described. A record in an array was created for each user session, while an extra array was inside it, preserving three variables: *time*, *direction* and *intensity* for each movement that was detected as shown in Figure 1(b).

## 4 Data Visualization

Data-Driven Documents (D3) is a novel representation-transparent concept for web-based visualizations. This JavaScript library assists users at bringing data to life using varied technologies such as HTML for page content, CSS for aesthetics, JavaScript for interaction, SVG for vector graphics and so on. D3's emphasis on web standards provides full capabilities of modern browsers while it combines powerful visualization components and a data-driven approach to a shared representation of the page called the *document object model* (DOM).

Highcharts is a charting library written in pure JavaScript which suggests an easy way of adding interactive charts to web applications. Many of these can be combined in one chart. The default framework implementation of Highcharts uses jQuery [2011], hence the only requirement for users is to load the jQuery library before Highcharts.

The first visualization, as shown in Figure 1(c), is a 2D scatterplot which displays the player's movement vs time (only o a zoomed portion of the whole graph). The x axis represents the time scale in seconds during which the tests take place while each label in y axis symbolizes each different user who participates in the test. Four different arrows imitate the movement of the human's head in two DOF. Furthermore an additional feature is displayed when the mouse is hovering an arrow, showing the respective time the movement occurred and the intensity which is based on how large was the difference between the previous and the current. Apart from those elements, a color fluctuation is also evident which serves as an intensity indicator for each movement (the closer to red color the arrow is, the higher the intensity of the movement). The full version of the web-based visualization can be found at: http://83.212.117.19/HeadPoseScatterplot/.

The second visualization consists of a 3D column diagram which illustrates the aggregation of all users' movements grouped by direction every two seconds as shown in Figure 1(d). The four different directions are imitated by four different colors. In one hand, x axis represents the time scale which is divided every two seconds until the end of the test. On the other hand, y axis displays the number of movements for all the users that take part in the tests. Furthermore, when hovering above a column, the number of the corresponding direction summary is displayed. Moreover, not so evenly distributed movements, see for example columns between 2-4 seconds in Figure 5, can lead into practical conclusions taking into account the nature of the test as well. The full version of the web-based visualization can be found at: http://83.212.117.19/HeadPose3D/.

## 5 Conclusions

A scheme for efficient and user friendly visualization, for the sake of improving both the understanding and analysis of data captured from a real time head pose estimation system, was presented. The selected system for head pose estimation was extended in an applicable way for collecting and ultimately storing the required data using a key-value style lightweight exchanging format, JSON. Finally two different visualizations derived from the data compilation stage that can be easily accessed on the web.

The intention of the authors was to provide the necessary visualizations in order for an objective evaluation task on those findings to follow.

## 6 Acknowledgments

The accomplishment of this poster was co-funded by the project "IKY Scholarships" from resources of the EP "Education and Lifelong Learning" of the European Social Fund (ESF) of ESPA, 2007-2013

## References

Fanelli, Gabriele, Dantone, Matthias, Gall, Juergen, Fossati, Andrea, and Van Gool, Luc. Random forests for real time 3D face analysis. *International Journal of Computer Vision*, 3, 101 (August 2012), 437-458.

Fanelli, Gabriele, Weise, Fanelli, Gall, Juergen, and Van Gool, Luc. Real Time Head Pose Estimation from Consumer Depth Cameras. In Heidelberg, Springer-Verlag Berlin, ed., Computer Vision and Pattern Recognition (CVPR). Springer Berlin Heidelberg, 2011.

jQuery. http://jquery.com/, Mar 2011.

Murphy-Chutorian, Erik and Trivedi, Mohan Manubhai. Head Pose Estimation in Computer Vision: A Survey. *Pattern Analysis and Machine Intelligence, IEEE Transactions on* , 31, 4 (April 2009), 607 - 626.