Kinect Annotation and Evaluation Tool Readme

Availability

Project homepage: http://www.multimedia-computing.de/wiki/Kinect_Annotation_and_Evaluation_Tool.
Source project: https://github.com/ChrislS/Kinect-Annotation-and-Evaluation-Tool .

Licenses:

To create the project, several libraries were used:

- Math.NET (unmodified, MIT/X11 license, http://numerics.mathdotnet.com/),
- MediaSlider (modified, CPOL (CodeProject Open License) , http://www.codeproject.com/KB/progress/MediaSlider.aspx),
- File sort algorithm (modified, public domain, http://beta.unclassified.de/code/dotnet/naturalsort/),
- OpenNI (unmodified, LGPL, http://www.openni.org/downloadfiles).

For the project source code itself, copyright 2011 Christoph Lassner, all rights reserved.

This project source code is released free under the terms of the GPL license version 3. You can find all applicable licenses after the installation in the installation sub-directory named "Licenses".

Prerequisites:

- .NET Framework v4.0 or higher,
- A computer with 2Gb Ram and 2.6Ghz dual core processor or better,
- Sufficient free hard disk space and a fast drive.

Installation:

- Install OpenNI version 1.3.2.1 or newer.
    1) Unplug the Kinect.
    2) Uninstall all possibly previously installed Kinect drivers and OpenNI versions (search in Software for "OpenNI ", "PrimeSense Sensor KinectMod" and "PrimeSense - NITE", in the Device Manager for "PrimeSense" and click "Uninstall Driver" for each of the sub-objects).
    3) Install OpenNI (v1.3.2.1 or newer).
    4) Install NITE (Use the following license key: **0KOIk2JeIBYClPWVnMoRKn5cdY4=**).
    5) Install SensorKinect and restart your computer (required!!!).
    6) Plug in the Kinect.
- Run the provided installation file (setup.exe) to install the program.
- Enjoy using it!

Documentation

The program is intended to be used for the following workflow:

1. (Select a project directory)
2. Record training or evaluation samples (with the Kinect, including 3D annotation data)
3. Review recorded samples, select frames for export
4. Export frame images and, if selected, annotation data with a selected configuration (e.g. background-separated)
5. (Train and export your own algorithm results to an xml-File)
6. See the algorithm results compared to Kinect annotations and evaluate the mistakes
7. (Add new features and contribute to the project).

It is possible to do step 2 to 5 several times with different configurations. On the following pages, you can find detailed information for each of these steps and the appendix contains details to the project file structure, xml-file format and others.

Step 1: Select a project directory

Start the Kinect Annotation and Evaluation Tool. The main window will look like below. Select a project directory by clicking on the button "Change".
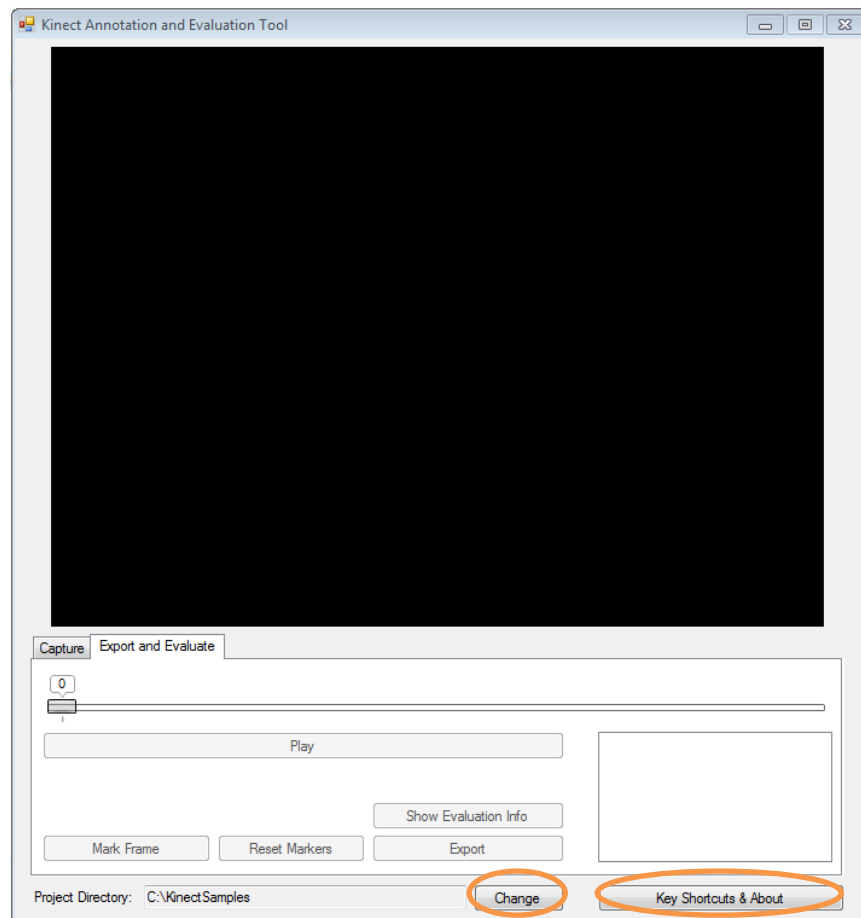
The project directory will be saved and automatically set at the next start of the program. It is possible to use the program for export and evaluation even if no Kinect is connected to the computer. To get information about the program and the key shortcuts, use the button "Key Shortcuts & About".

Step 2: Record training or evaluation samples

After starting the program, click on the tab "Capture". The Kinect must be connected to your computer to be able to select this tab, otherwise you will be prompted with a hint accordingly. It is not possible for the program to detect the device if you attach it to the computer after the program has been started, so if necessary quit the program, attach the device and then start it again.
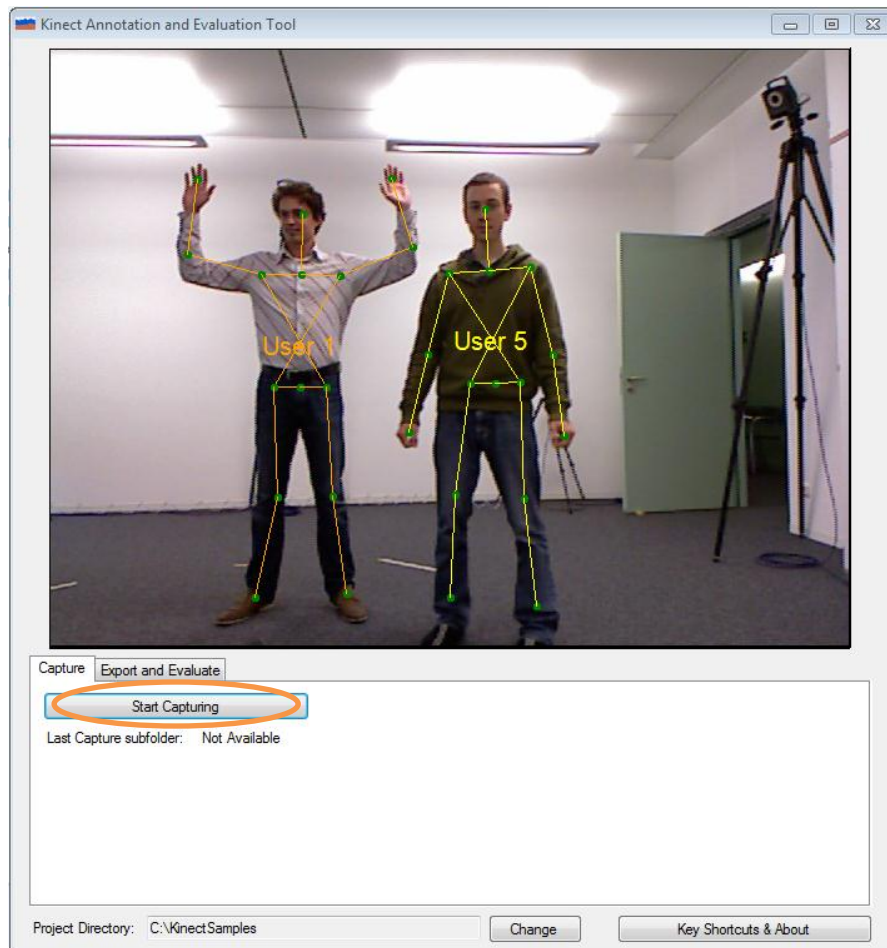
The black area should now display the live video stream from the connected Kinect device. You can experiment with the key shortcuts to modify the displayed video stream and make any adjustments to the environmental settings for the recording. You can look the key shortcuts up by pressing "Key Shortcuts & About".

When you are ready to start capturing, click the button "Start Capturing" or press "Enter". The live display of the video will change to black and only the skeletons of tracked users will be displayed. This is intended to save performance to capture as fast as many frames as possible. However, if you need specific visual feedback for the recording, you can change the live display mode at any time by using the corresponding key shortcut.

Once you finished recording, click the button "Stop Capturing" or press "Enter" again. The line "Last Capture subfolder" displays the name of the subfolder in the project directory where the recorded data was saved.

It is possible to start capturing the next take right away. Once you finished recording, select the tab "Export and Evaluate".

Step 3: Review recorded samples, select frames for export

The list on the right hand side of the "Export and Evaluate"-Tab shows all detected takes in the selected project directory. It is possible to view the content of a take simply by clicking on the name in the list. You can skim through the video using the frame indicator bar and use all key shortcuts similar to the live view mode. In playback mode, the frames are played at the same frame rate as they were recorded. If you encounter changing playback rates, this is due to the lack of processing power while recording.
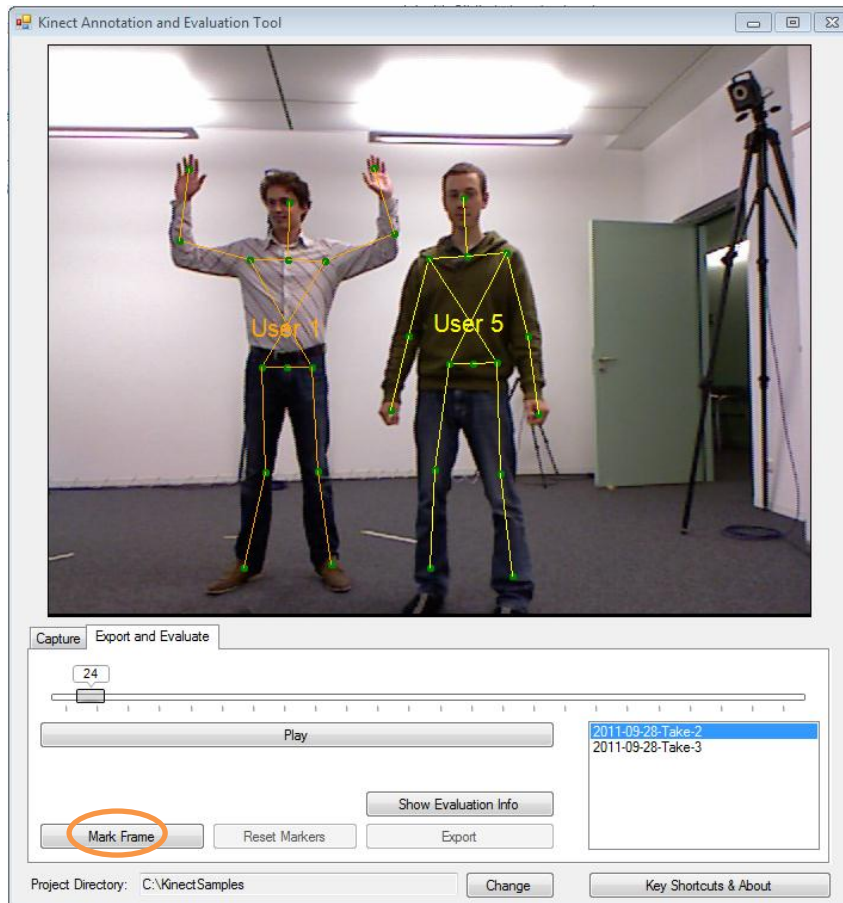

Once you identified a series of frames you want to export, go to the first frame and select "Mark Frame". Go to the last frame and select "Mark Frame" again. If you want to extend your selection, you can go to an earlier or later frame than the selected and use the "Mark Frame" button again. Alternatively, use the "Reset Markers" button to remove the markers and restart the selection.

The joint indicators are green if the Kinect assumes a confidence value > 0.5 for a skeleton joint, red otherwise. In the current OpenNI implementation, only two values (0f and 1f) are used.

Step 4: Export frame images

Once you have finished your selection of important frames as described for Step 3, click the "Export" button. A dialog window will be displayed like the following:
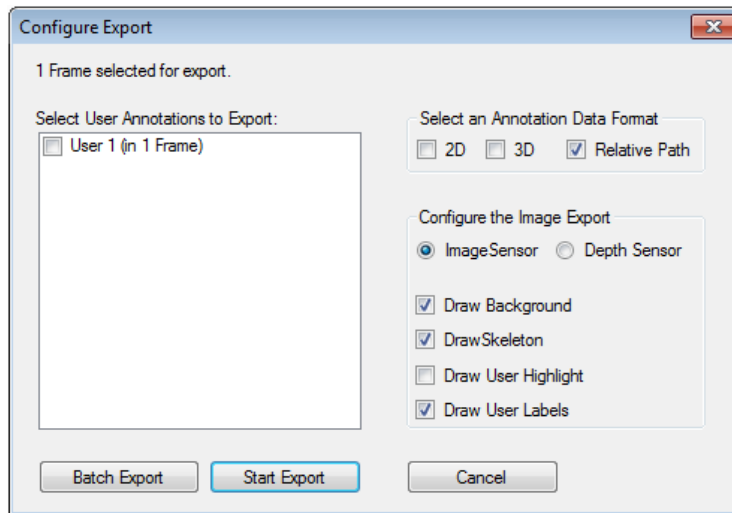
If you plan on using annotations for the exported images, select the users you want to export the annotations for. It is then necessary to also select one or more annotation data formats. The "Relative Path" option specifies whether the filename of the image file the annotation is for in the .xml annotation file is relative to the .xml file or an absolute name.

The "Configure the Image Export" section offers various options to modify the image content. If you choose to export the depth sensor data, the <u>fake </u>depth images will be exported as displayed in the video live view.

The button "Batch Export" allows to export images and annotations as selected for the following configurations with one click (the configuration in the "Configure the Image Export" section will be ignored):

1. Plain image sensor images
2. Plain image sensor images with skeleton
3. Background separated image sensor images

Click "Batch Export" or "Start Export" to start an export with the corresponding configuration. The images will be exported to single image files and stored in a folder called "Export-[number]" subdirectory of the project. The image files have filenames starting with their number and a suffix (see the appendix for an explanation of suffixes). An xml file will be exported with annotations and image names.

## Step 5: Use the exported data

The format of the .xml-files in the export directories is described in the appendix.

Use the exported data to train your algorithms and save the results in the same format in the export-directory in a .xml file with the prefix "Results-" (case sensitive!). See the appendix for a more detailed explanation of the directory structure.

Step 6: Evaluate

In the "Export and Evaluate" tab, select the take you used your algorithm with, to evaluate it. Click the "Show Evaluation Info" button to open the evaluation window.

If you saved your algorithm results to an .xml file with the prefix "Results-" as described in Step 5, it should be listed in the "Configure Data Source" section in the respective export folder. Click on it and select a user in the "User Results" section (your algorithms results) and one in the "User Ground Truth" section to start calculating the errors.
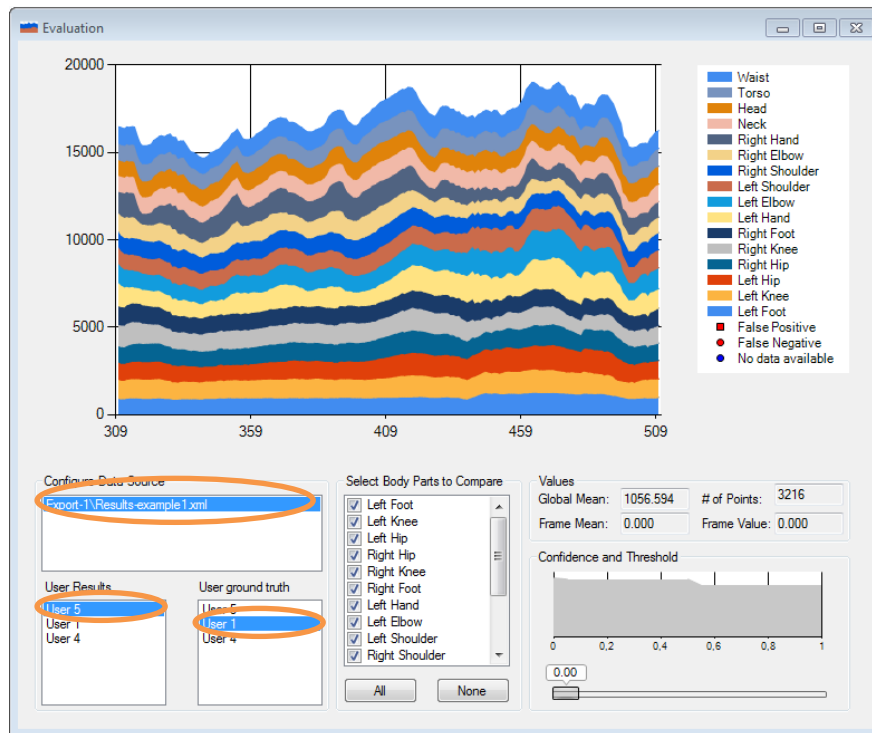
Use the section "Select Body Parts to Compare" to constrain the comparison to specific body joints. The section "Confidence and Threshold" displays a cumulative confidence diagram for the selected joint points. By modifying the threshold value it is possible to restrict the comparison only on detected joints, OpenNI gave a confidence value better than the selected. However, as currently this value, despite it is a float-value, gets only the two values 0f and 1f, the benefit of this feature is limited. With a future OpenNI release, this might change.

By using the frame indicator in the main window, it is possible to go to the evaluated frames and see the frame error for the frame as well as to see the joint positions estimated by the algorithm in the video, as well as the ones estimated by the Kinect. In the error diagram, the current frame is indicated by a black vertical line.

Step 7: Contribute

The program is hosted on GitHub to make collaboration and integration of extensions and modifications easy. If you have built your own new feature or made an improvement, we will be happy to integrate it, to share it with all others who are using the program.

If you have questions or any other requests, feel free to contact me at Christoph.Lassner@googlemail.com .

Appendix

Joints and their representations

The joint translation is located in the file "JointDictionary.cs" in the region "Joint Conversion" and can be modified easily. As the application does not use the "GlobalMirror" option, the OpenNI skeleton joints and the real names are right-left-switched.

Real name and Evaluation window	OpenNI SkeletonJoint	.xml Export ID
Right Foot	LeftFoot	0
Right Knee	LeftKnee	1
Right Hip	LeftHip	2
Left Hip	RightHip	3
Left Knee	RightKnee	4
Left Foot	RightFoot	5
Right Hand	LeftHand	6
Right Elbow	LeftElbow	7
Right Shoulder	LeftShoulder	8
Left Shoulder	RightShoulder	9
Left Elbow	RightElbow	10
Left Hand	RightHand	11
Neck	Neck	12
Head	Head	13
Torso	Torso	15
Waist	Waist	14

The Waist joint is not yet supported by the OpenNI framework. Because the joint information was needed to evaluate a specific algorithm, the position is a synthetic one, calculated as the center of both hip-joints.

.Xml file format

```
<?xml version="1.0"?>
<annotationlist Is3DData="[True|False]">
        (For each annotated image:)
        <annotation>
                (either:)
                <image>
                        <name>(Image name. Must be a number, optionally followed by "-" and
                                prefix, then "." and format specifier.)</name>
                </image>
                (or:)
                <frame-number>#</frame-number>
                (end either/or)
                (For each user in the image:)
                <annorect UserID="#">
                        <annopoints>
                                (For each skeleton joint:)
                                <point>
                                        <id>#</id>
                                        <confidence>#</confidence>
                                        <x>#</x>
                                        <y>#</y>
                                        (Iff 3D data:)
                                        <z>#</z>
                                </point>
                        </annopoints>
                </annorect>
        </annotation>
</annotationlist>
```

File structure

```
<Project directory>
        <Take directory "[date]-Take-[number]">
                "KinectRawData.oni" (video and depth sensor recording)
                "KinectUserData.raw" (User mask recording. Binary file with 640*480 ushort values
                                per frame.)
                "KinectUserPosition.xml" (xml annotation file. Format as specified before.)
                <Export directories "Export-[number]">
                        <Image files "[number]-[suffix].png">
                        <Annotation files "Export[2|3]D-[suffix].xml" Format as specified before.>
                        <Your own result files "Results-[anything].xml" Format as specified before.>
```

File suffixes

The plain image sensor images are assumes as "normal" export mode. These images and annotation files do not get any suffixes. All files exported in other modes get suffixes as specified below (take into account the similarity to the key shortcuts to enable the respective mode):

Mode	Letter added to file suffix
Skeleton Drawn	s
No sensor data	a
Depth sensor instead of image sensor selected	d
Background omitted	b
Highlighted user areas	h
User labels	l

Key Shortcuts

d: (depth) use depth sensor data as information source

i: (image) use image sensor data as information source

a: (any) toggle drawing of any sensor information

l: (label) toggle user label display

s: (skeleton) toggle skeleton display for tracked users

b: (background) toggle background display

h: (highlight) toggle highlighting users

'enter': Start/Stop recording if button is available

m: (mark) frame as begin or end

r: (reset) frame markers

e: (export)

p: (play/pause)