iKarate: Improving Karate Kata

By: Bassel Emad
Omar Atef
Yehya Shams
Ahmed El-Kerdany

Supervised By: Dr. Ayman Ezzat
Dr. Ayman Nabil

Teaching Assistant: Eng. Nada Ayman

In Collaboration With:

8/10/2019

INTRODUCTION 1/2

Martial art that can be practiced using hands and feet to deliver and block strikes.

Kata

- Sequence of movements.
- Offensive and defensive postures.

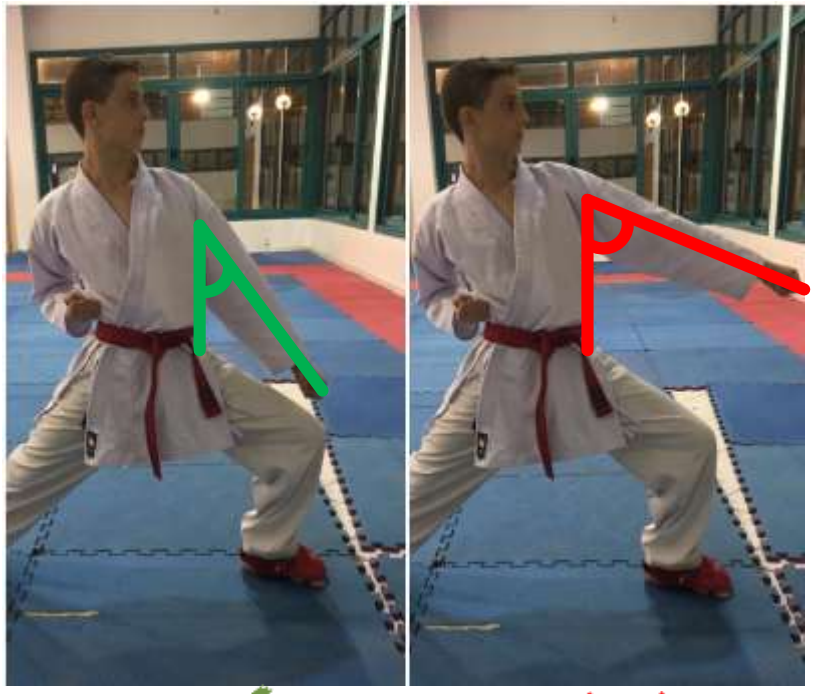
Heian Shodan Kata 1

INTRODUCTION 2/2

- **Common Mistakes:**
  - Wrong Angles.
  - Joints Positioning.
  - Arms, Legs, And Fists.

Age Uke

Gedan Barai



RELATED WORK 1/3
BEST KINECT CLASSIFIERS

- Analysis of several machine learning techniques (**Accuracy** - Precision - **Time**).
- 6 Data sets used from Kinect 1 and Kinect 2 tested with **41 classifiers**.
- Computation time is higher for the Kinect 2 datasets than for Kinect 1.
- Best overall performance is **Multilayer Perceptron** (Accuracy **99.0874%**).

TABLE II: Best Classifier Results on Each of the Six Datasets Collected with Kinect 1 and Kinect 2

Dataset	K1 20M	K1 30S	K2E 20M	K2E 30S	K2R 20M	K2R 30S
Best 11 classifiers	>89%	>89%	>95%	>97%	>96%	>96%
Best Classifier (BC)	Simple Logistic	Multilayer Perceptron	Random Forest	Multilayer Perceptron	Random Forest	Random Forest
BC Precision	98.30%	97.70%	98.90%	99.10%	98.90%	99.00%
BC Accuracy	98.20%	97.60%	98.82%	99.08%	98.82%	98.95%
BC Time (sec)	5.54	33.47	0.88	65.93	0.83	1.15

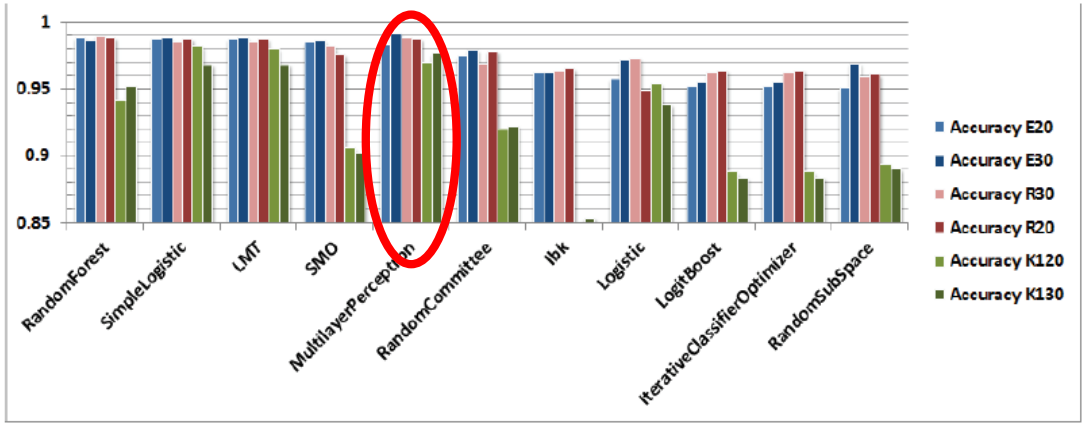


Fig. 1: Accuracy for the best 11 classifiers on the 6 datasets collected with Kinect 1: K120M (K120), K130S (K130) and Kinect 2: K2E20M (E20), K2E30S (E30), K2R20M (R20), K2R30S (R30).

Calin, A. D. (2016). **Variation of pose and gesture recognition accuracy using two kinect versions**. 2016 International Symposium on INnovations in Intelligent SysTems and Applications (INISTA). doi: 10.1109/inista.2016.7571858

RELATED WORK 2/3
REAL-TIME RECOGNITION

- Mechanism for **real-time recognition** of 3D gestures.
- Mahalanobis distance and **Dynamic time warping**.
- Similarity between unknown and known gesture is found from the sum of DTW and Mahalanobis distances.
- Total average recognition rate of **84.5%**.

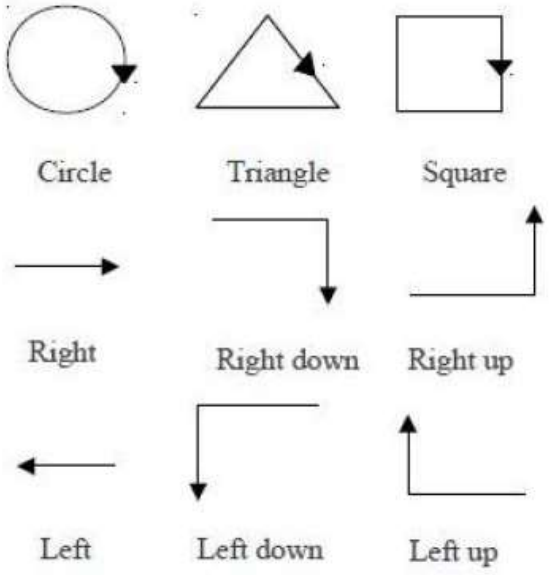


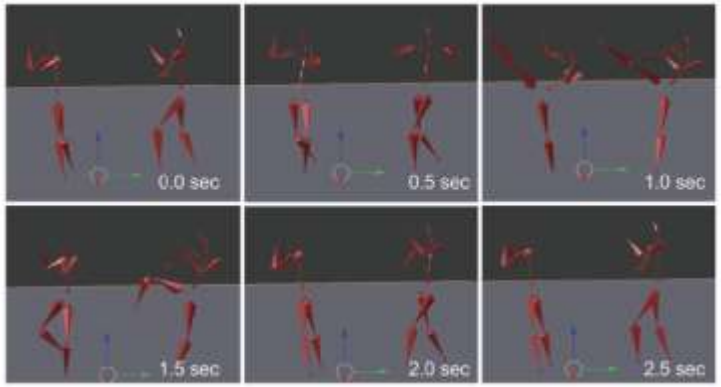
Fig. 4. Gesture used in the experiment

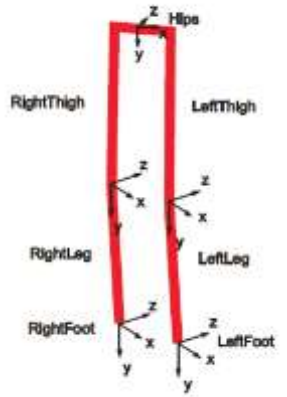
Gesture	Observed Recognition Results			
	Total	Correct Det.	False Det.	Rec. Rate
Circle	50	44	6	0.88
Square	50	44	7	0.86
Triangle	50	41	9	0.82
Left	50	41	9	0.82
Left Up	50	42	8	0.84
Left Down	50	45	5	0.90
Right	50	42	8	0.84
Right Up	50	44	6	0.88
Right Down	50	44	6	0.88

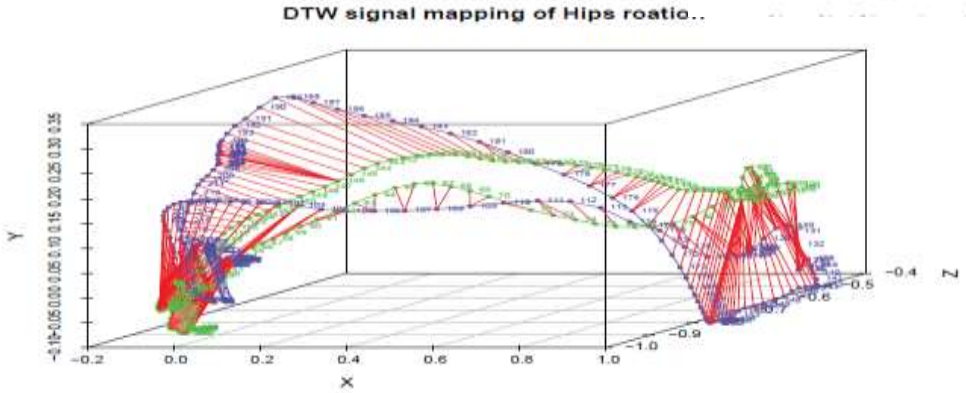
Janaki, V. M. S., Babu, S., & Sreekanth, S. S. (2013). **Real time recognition of 3D gestures in mobile devices**. 2013 IEEE Recent Advances in Intelligent Computational Systems (RAICS). doi: 10.1109/raics.2013.6745463

RELATED WORK 3/3
BODY SPEED AND PROPORTIONS

- Actions could be performed in **different speeds** and **different body proportions**.
- Using Shadow toolkit.
- Angle **normalization** algorithm And **DTW** alignment algorithm.
- The DTW signal mapping plots are even more informative when one can rotate them and change number of points to display in order to see the kinematic progress.





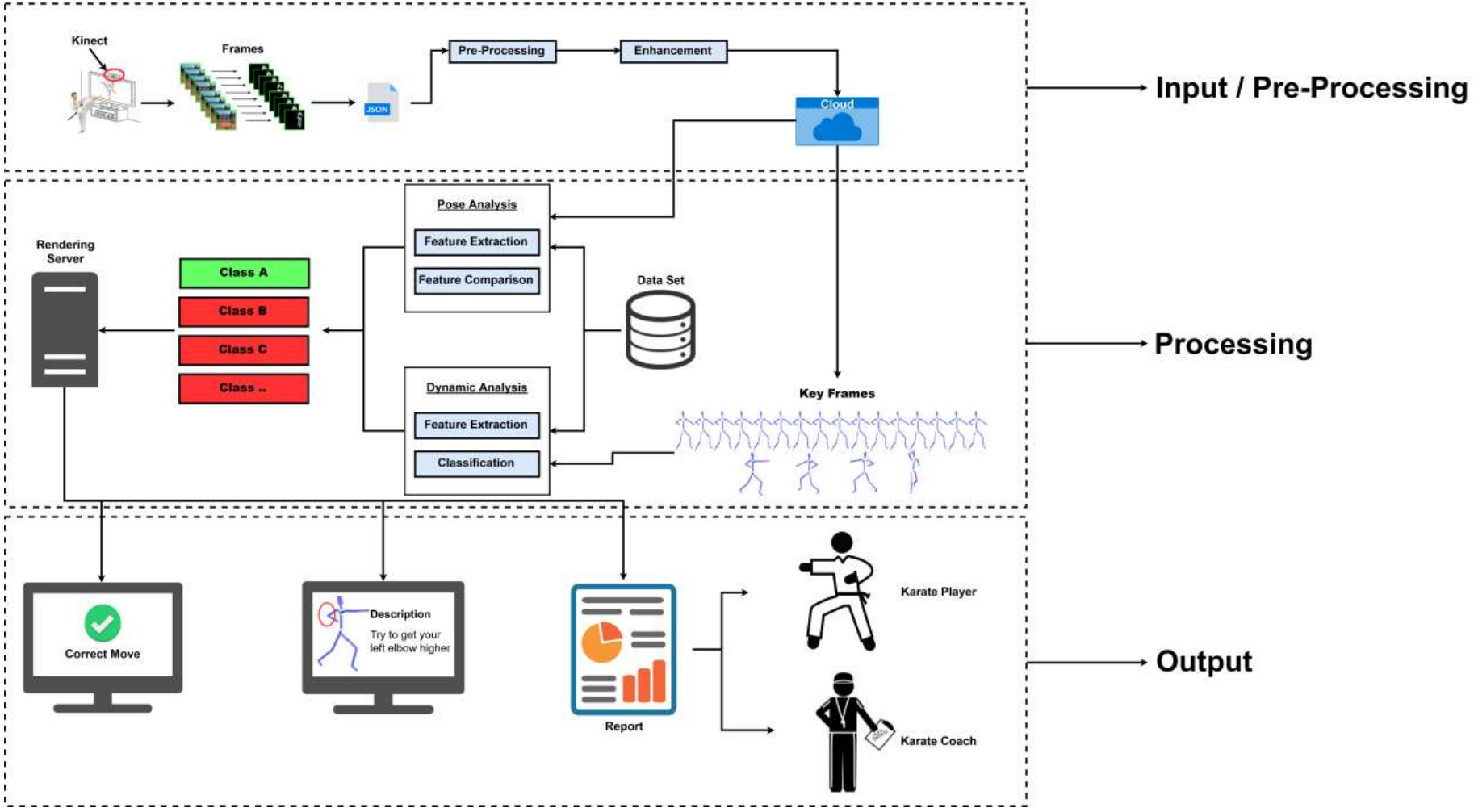


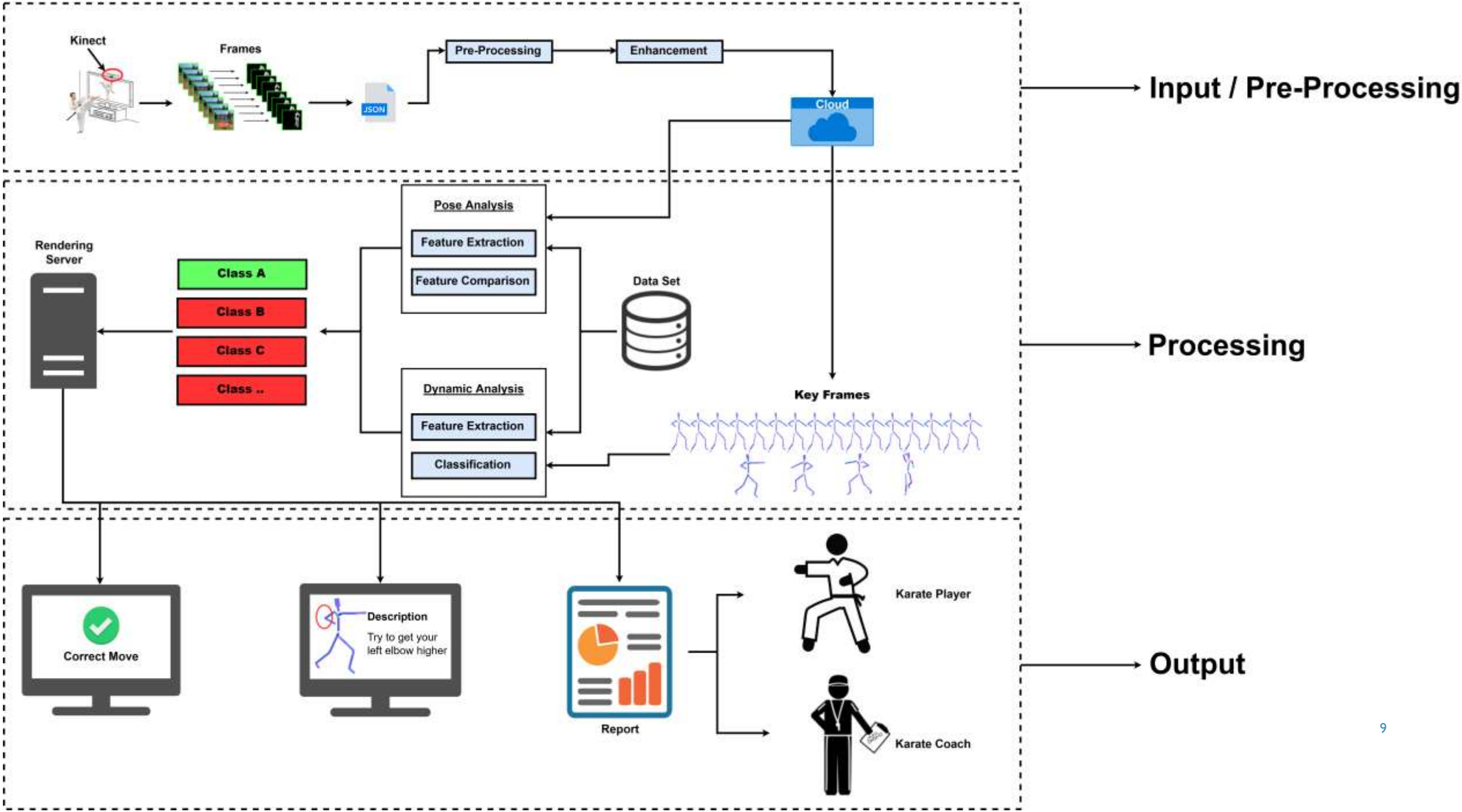
Tomasz Hachaj, Marek R. Ogiela, Marcin Piekarczyk, and Katarzyna Koptyra. **Advanced human motion analysis and visualization: comparison of mawashi-geri kick of two elite karate athletes**. IEEE Symposium Series on Computational Intelligence (SSCI) IEEE Symposium Series on Computational Intelligence (SSCI) (2017). DOI: 10.1109/SSCI.2017.8285269

PROBLEM STATEMENTS

Enhancing the Classification Accuracy And Time of karate kata, providing Real-Time Feedback and taking into consideration the Different Speed And Body Proportions.

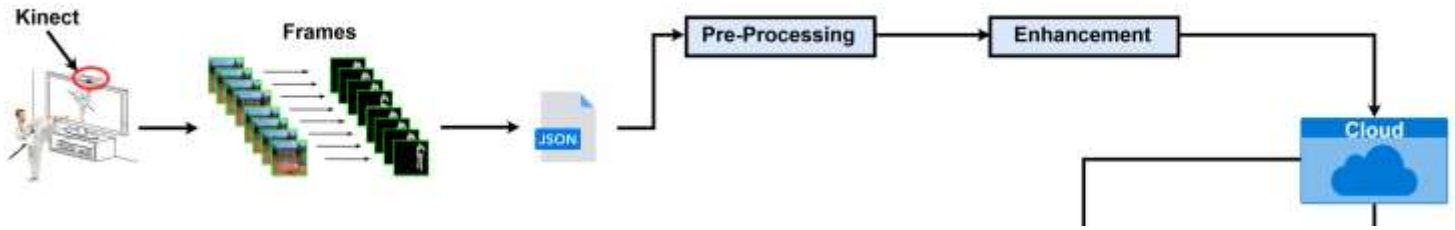
SYSTEM OVERVIEW




Kinect
Frames
JSON
Pre-Processing
Enhancement
Cloud
Input / Pre-Processing
Pose Analysis
Feature Extraction
Feature Comparison
Rendering Server
Class A
Class B
Class C
Class ..
Data Set
Dynamic Analysis
Feature Extraction
Classification
Key Frames
Processing
Correct Move
Description
Try to get your left elbow higher
Report
Karate Player
Karate Coach
Output

PRE-PROCESSING

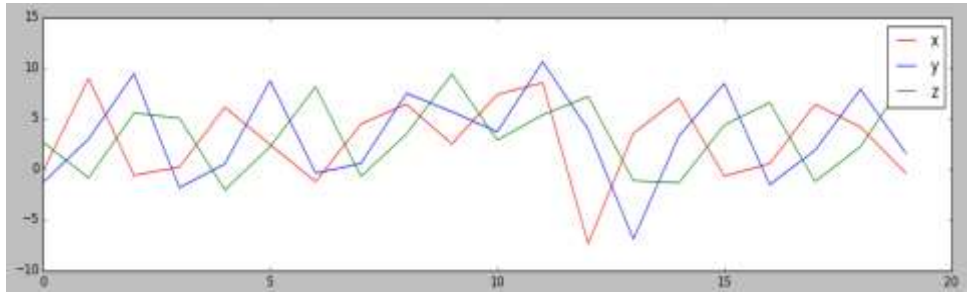
- Get Xbox Kinect(s) Readings.
- Proposed Filtering Techniques:
  - Low-Pass.
  - Hamming.
  - Kalman.



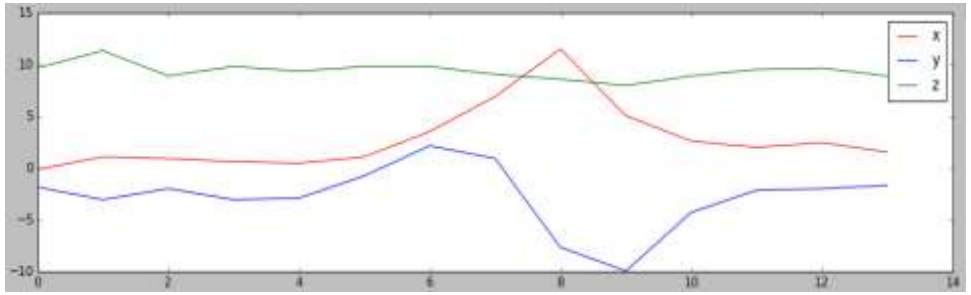
- Outliers Segmentation:
  - Isolation Tree.

- Normalization.

- Signal Interpolation And Extrapolation.



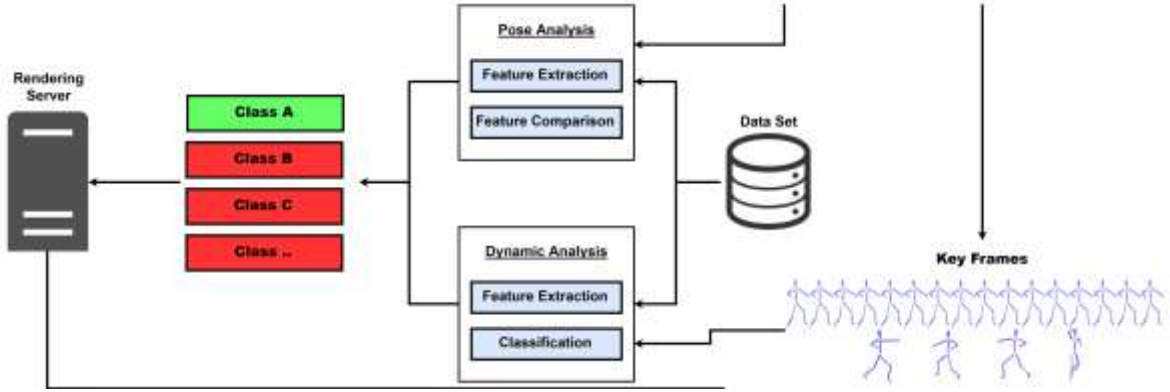
Before



After

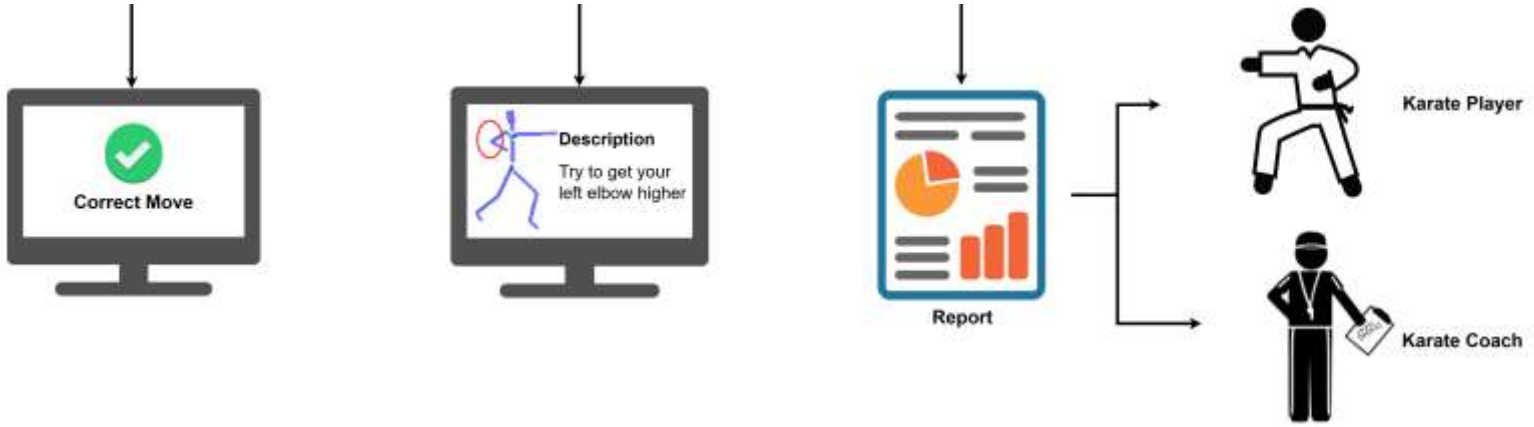
PROCESSING

- **Proposed Classification Algorithms:**
  - Fast DTW (Dynamic Time Warping).
  - SVM (Support Vector Machine).
  - R-CNN (Region-Based Convolutional Neural Networks).

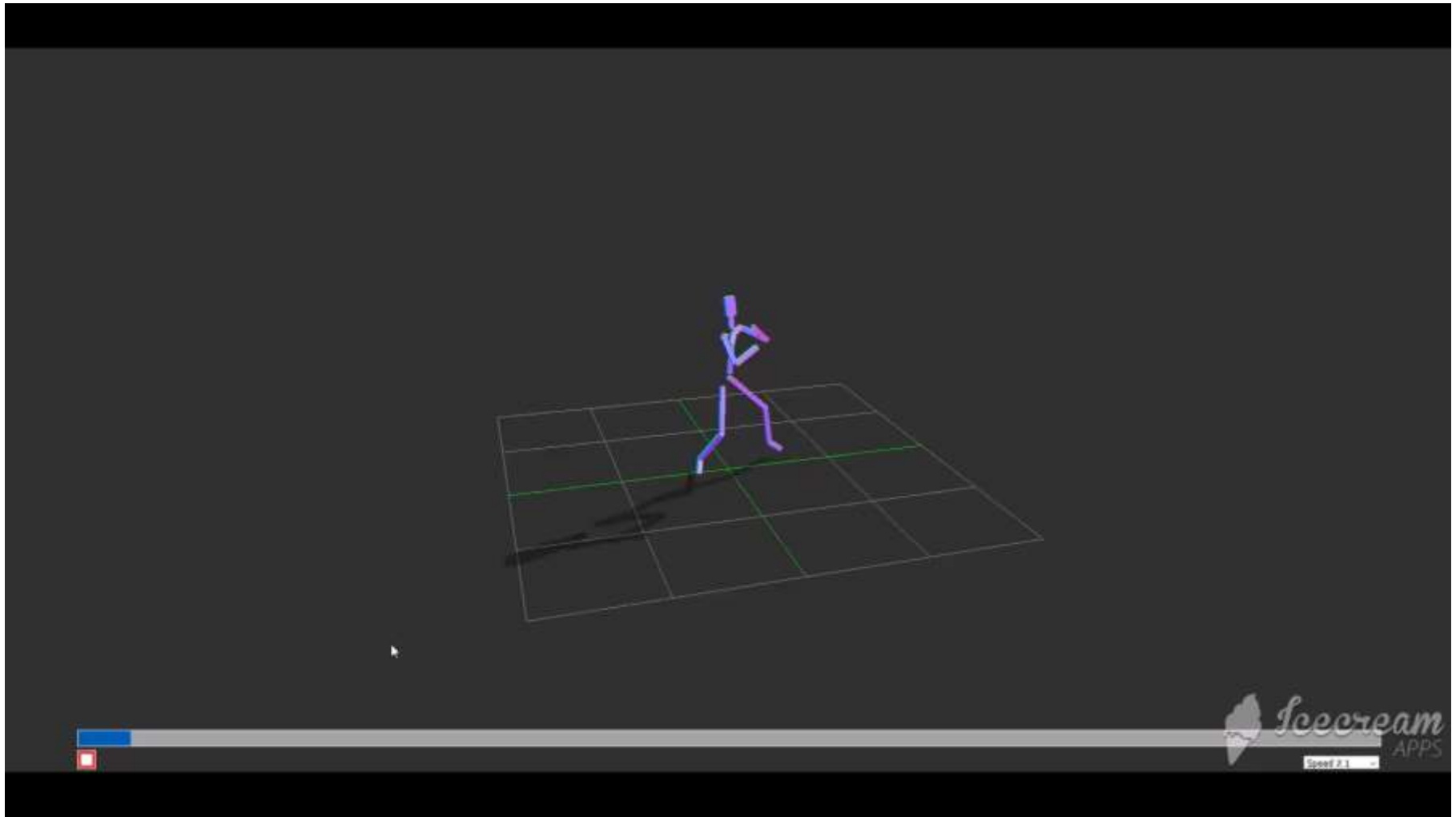


EXPECTED RESULTS

- **Reports:**
  - Automatic Scoring.
  - Movement Name.
  - Movement Rate.
  - Movement Duration.
  - Real-Time Feedback.



DEMO



DATA SET 1/2

Asking for help in Karate and GDL Papers

Inbox ×

Omar Omar Atef Hashem Ahmed <omar1603910@miuegypt.edu.eg> Wed, Aug 21, 10:04 PM

to tomekhachaj, mogiela, marcin.piekarczyk, kkoptyra

Hello,

My name is Omar Atef, I am a Computer Science Student, who studies in Egypt at Misr International University (MIU). Currently a senior and working on my Graduation Project with 3 other friends. Our project subject is about Karate, and developing a system using Kinect that could help learners get more professional and adjusting their skill motion. This system could be able to recognize the performed movement and give a fair feedback about how correct it was and giving hints to improve his skills.

I am sending this mail asking for your help, Since i read about your published papers in IEEE and knew that you are an expert in this field. So if you could allow me to ask you some questions and giving me some hints and guidance to complete my project I would really appreciate it.

best regards.

DATA SET 2/2

tomekhachaj Sat, Aug 24, 7:01 PM

to mogiela@agh.edu.pl, marcin.piekarczyk@up.krakow.pl, me, kkoptyra@agh.edu.pl

Hello,
Main issue with karate is to gather high - quality data. You can do it with Kinect, however it might be difficult to precisly compare to motion capture recordings of two persons. If you do not have access to motion capture hardware, you can use our data for tests. You can download data from here:
http://gdl.org.pl/
to load and process / compare data I recommend this R language package:
https://github.com/browarsoftware/RMoCap
I recommand to work on bvh files. There is a documentation with examples and full source code.
If you would like to use any of those resource plese cite those papers:

Tomasz Hachaj, Katarzyna Koptyra, Marek R. Ogiela, Averaging of motion capture recordings for movements' templates generation, Multimedia Tools and Applications, December 2018, Volume 77, Issue 23, pp 30353–30380, DOI: 10.1007/s11042-018-6137-8

Tomasz Hachaj, Marcin Piekarczyk, Marek R. Ogiela, Human actions analysis: templates generation, matching and visualization applied to motion capture of highly-skilled karate athletes, Sensors (Basel). 2017 Nov 10;17(11). pii: E2590. doi: 10.3390/s17112590

Tomasz Hachaj, Marek R. Ogiela, RMoCap: An R package for processing and kinematic analyzing motion capture data, Multimedia Systems, In press

Best regards,
Tomasz Hachaj

•••

POTENTIAL CUSTOMER

Any questions


THANK
YOU!

Appendix

DEFINITIONS 1/3

- **Kihon:** Kihon is a Japanese term meaning “basics” or “fundamentals”. The term is used to refer to the basic techniques that are taught and practiced as the foundation of most Japanese martial arts.

- **Kumite:** Kumite literally translated means "grappling hands" and is one of the three main sections of karate training, along with kata and kihon. Kumite is the part of karate in which a person trains against an adversary, using the techniques learned from the kihon and kata.

- **Heian Shodan:** Heian shodan is a shorin kata containing 21 movements with 2 kiai points. It is the first kata you learn when joining Shotokan karate. In Japanese, Heian means “peaceful mind” and shodan means “first level“.

- **Martial Art:** Martial arts are codified systems and traditions of combat practiced for a number of reasons such as self-defense; military and law enforcement applications; competition; physical, mental and spiritual development; and entertainment or the preservation of a nation's intangible cultural heritage.

DEFINITIONS 2/3

- **Shotokan:** Shōtōkan is a style of karate, developed from various martial arts by Gichin Funakoshi and his son Gigo Funakoshi.

- **Kiai:** Kiai is a Japanese term used in martial arts for the short shout uttered when performing an attacking move. Traditional Japanese dojo generally use single syllables beginning with a vowel.

- **Dojo:** A dōjō is a hall or space for immersive learning or meditation. This is traditionally in the field of martial arts, but has been seen increasingly in other fields, such as meditation and software development. The term literally means "place of the Way" in Japanese.

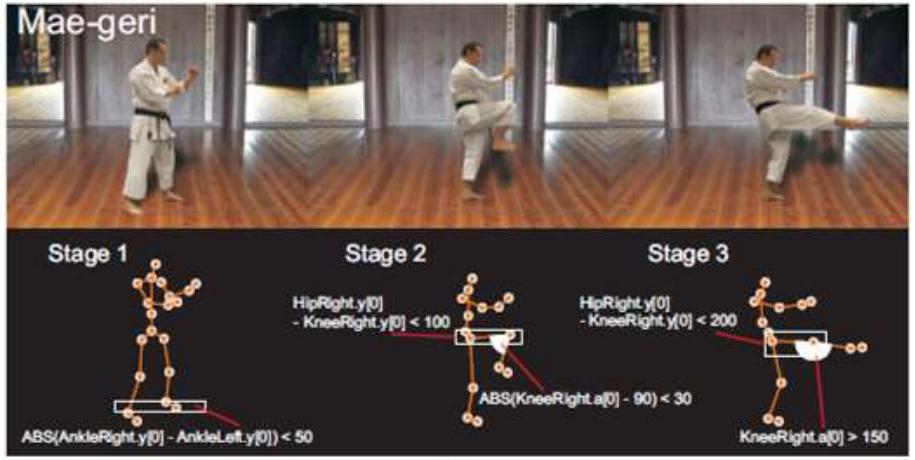
- **Shorin Ryu:** Shōrin-ryū is one of the major modern Okinawan martial arts and is one of the oldest styles of karate. It was named by Choshin Chibana in 1933, but the system itself is much older.

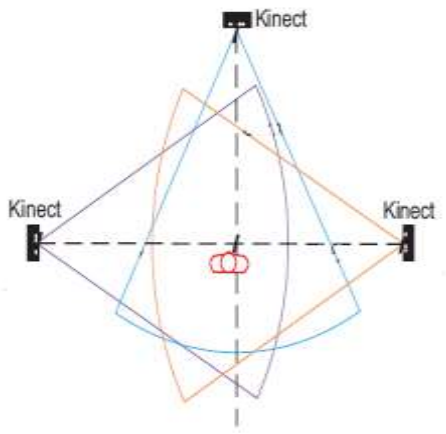
DEFINITIONS 3/3

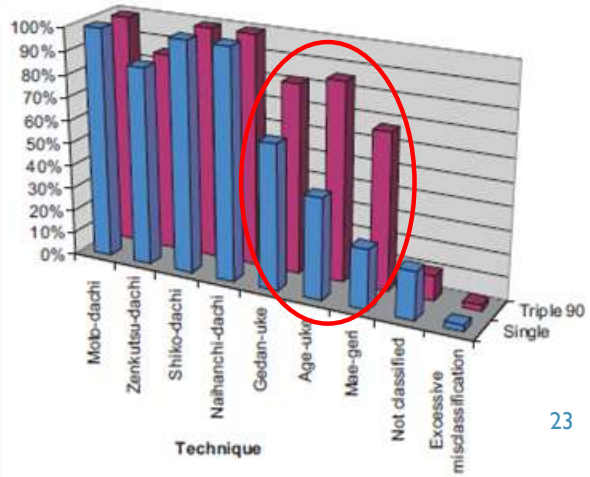
- **Okinawa:** Okinawa is a Japanese island in the East China Sea. In Naha city, Shuri Castle is the rebuilt palace of the Ryukyu Kingdom. One of several remaining Ryukyuan fortresses on Okinawa from the Gusuku period, it features the ornate gate of Shureimon. The Okinawa Prefectural Museum has exhibitions on Okinawa's natural and cultural heritage, plus a collection of fine art. Kokusai Street is lined with shops and restaurants.

RELATED WORK

- Comparison of motion detection effectiveness using **one** or **3 Kinects** with **different angle rotations**.
- Gesture Description Language (**GDL**) Classifier.
- Using the 90 and -90 degree will Improve non-classified movements with “48%”.
- Gedan-uke, Age-uke and Mae-geri accuracy were improved with 3 Kinects.



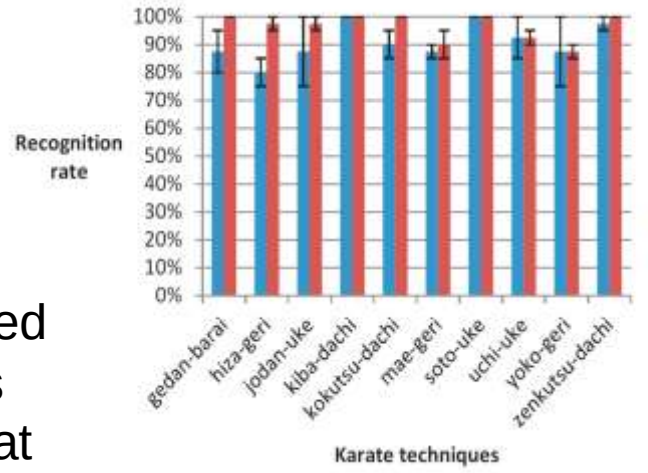




Tomasz Hachaj, Marek R. Ogiela, and Marcin Piekarczyk. **Dependence of Kinect sensors number and position on gestures recognition with Gesture Description Language semantic classifier**. Conference on Computer Science and Information Systems (2013) pp. 571–575

RELATED WORK

- Multidimensional time-varying signal, Actions might be similar to each other (False Positive Errors).
- GDL & **Reversed-GDL**.
- Using **Kinect 1** and **Kinect 2**.
- Ability to classify data from unsegmented recording that contains multiple actions and distinguish between key frames that belongs to different classes.



$t_i, t_{i+1}, \dots, t_j$ (i<j).

$$F_{[t_i,t_j]} = [\overline{f_{ti}}, \dots, \overline{f_{tj}}] \quad (1)$$

And $\overline{f_{ta}} \in \mathbb{R}^m$ where $a \in [i, j]$.

A signal sample belongs to cluster $C_b(\bar{\mu}, \bar{\sigma}, \bar{\varepsilon})$ when:

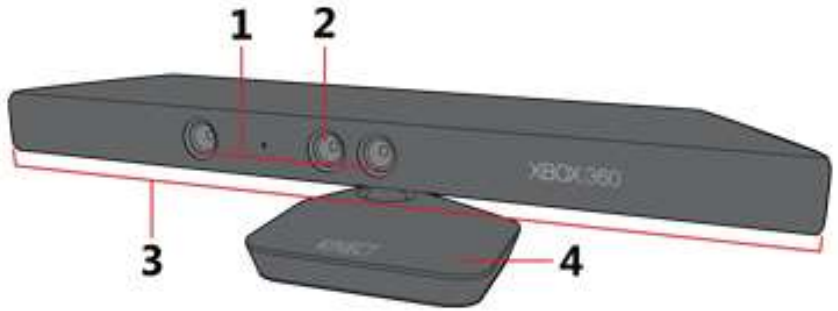
$$\overline{f_{ta}} \subset C_b \Leftrightarrow |\overline{f_{ta}} - \bar{\sigma}| < \bar{\mu} + \bar{\varepsilon} \quad (2)$$

$$GDL\left(F_{[t_i,t_j]}\right) = C_c \Leftrightarrow \begin{cases} \overline{f_{tl1}} \subset C_{c1} \wedge \dots \wedge \overline{f_{tln}} \subset C_{cn}, tl1 < \dots < tln \\ tj - tln < tcn \wedge \dots \wedge tl_1 - tl_2 < tc1 \end{cases} \quad (3)$$

T. Hachaj, M. R. Ogiela and K. Koptyra, **Effectiveness Comparison of Kinect and Kinect 2 for Recognition of Oyama Karate Techniques**, 2015 18th International Conference on Network-Based Information Systems, Taipei, 2015, pp. 332-337.doi: 10.1109/NBiS.2015.51

KINECT

1) 3D Depth sensor (IR Emitter + IR Camera/Depth Sensor).
2) RGB camera (Color Sensor).
3) Microphone array.
4) Tilt motor (For detecting the floor and players in the play space).

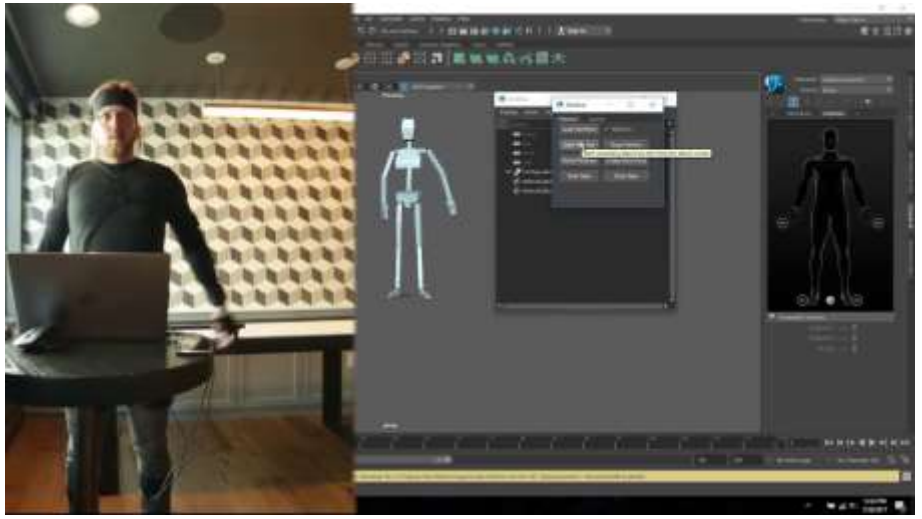


DIFFERENCE BETWEEN KINECT V1 AND V2

Feature	Kinect for Windows 1	Kinect for Windows 2
Color Camera	640 x 480 @30 fps	1920 x 1080 @30 fps
Depth Camera	320 x 240	512 x 424
Max Depth Distance	~4.5 M	~4.5 M
Min Depth Distance	40 cm in near mode	50 cm
Horizontal Field of View	57 degrees	70 degrees
Vertical Field of View	43 degrees	60 degrees
Tilt Motor	yes	no
Skeleton Joints Defined	20 joints	26 joints
Full Skeletons Tracked	2	6
USB Standard	2.0	3.0
Supported OS	Win 7, Win 8	Win 8
Price	$299	TBD

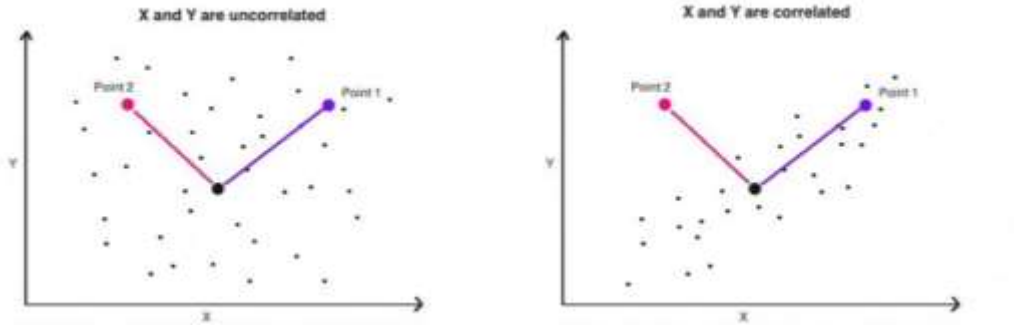
SHADOW MOTION CAPTURE (MOCAP)

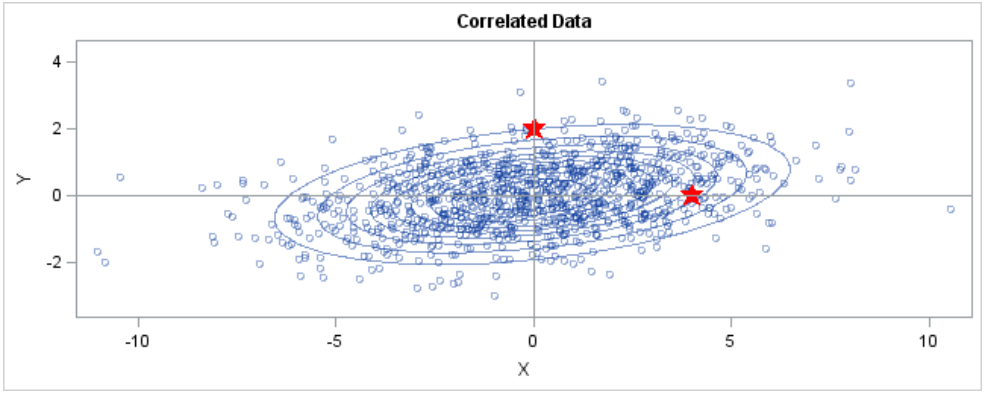
- The Shadow system uses the highest quality MEMS inertial components to provide low noise and high resolution output data for smooth and accurate orientation measurements.
- 17 precisions inertial measurement units.
- 3 Axis: accelerometer, gyroscope and magnetometer.
- 2 pairs of pressure insoles.



MAHALANOBIS DISTANCE

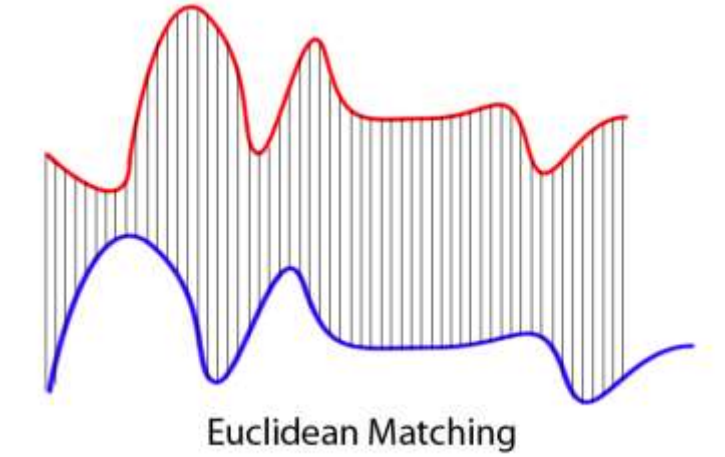
- Mahalanobis distance is the distance between two points in multivariate space.
- Mahalanobis distance measures distance relative to the centroid, a base or central point which can be thought of as an overall mean for multivariate data.
- The most common use for the Mahalanobis distance is to find multivariate outliers, which indicates unusual combinations of two or more variables.



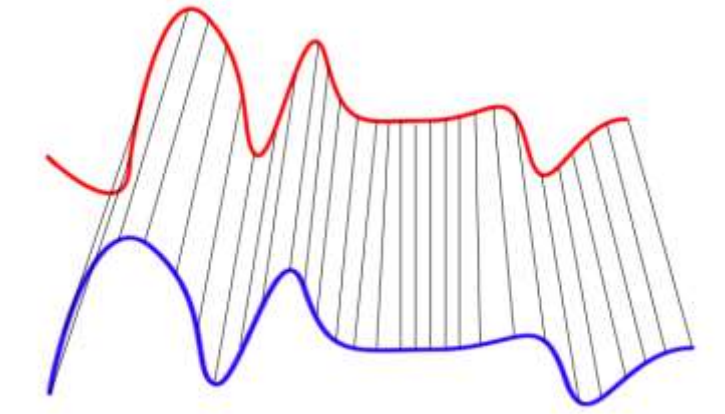


DYNAMIC TIME WARPING (DTW)

- Open-Source Library.
- Great for time variant data and speech recognition.



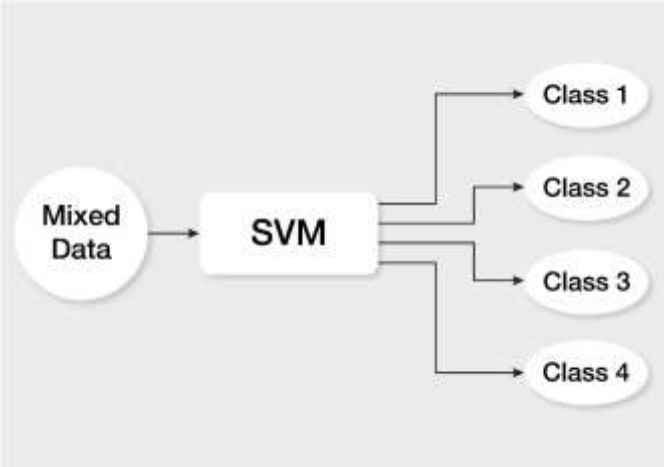
Euclidean Matching

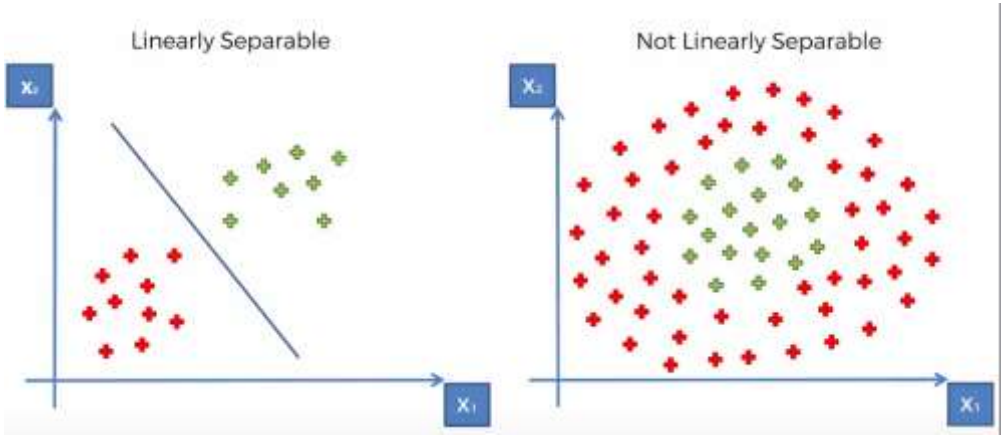


Dynamic Time Warping Matching

SUPPORT VECTOR MACHINE (SVM)

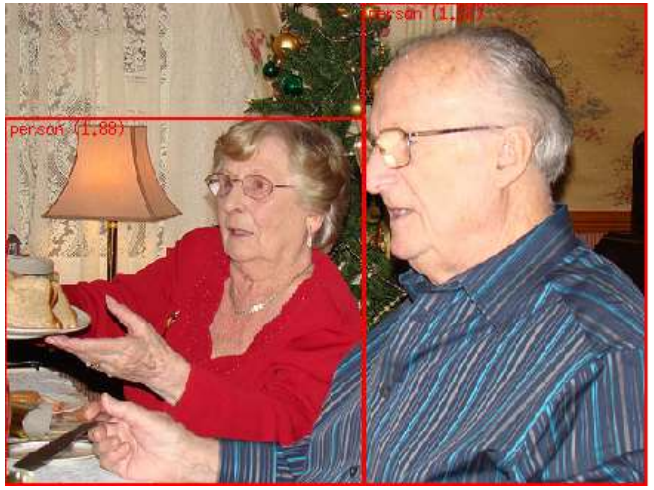
- SVMs are supervised learning models with associated learning algorithms.
- An SVM model is a representation of the data as points in space, mapped so that the data of the separate categories are divided by a clear gap.
- When data are unlabeled, supervised learning is not possible (Nonprobabilistic Binary Linear Classifier).





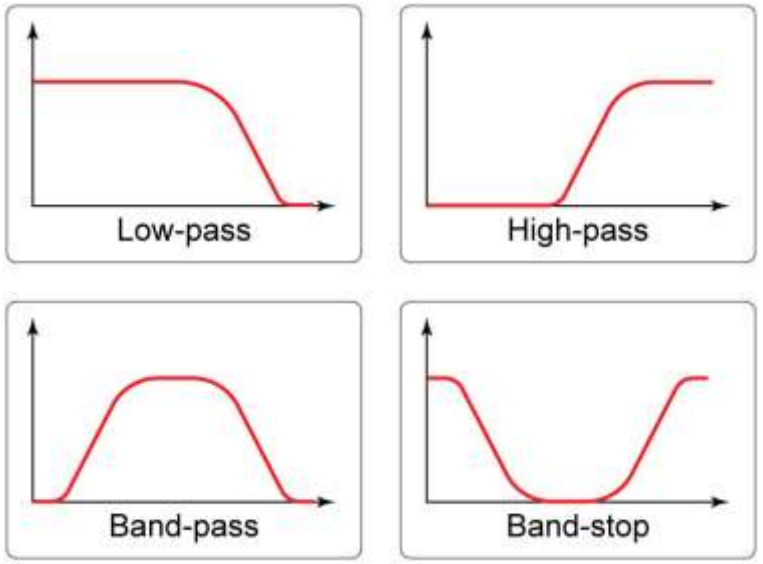
REGION-BASED CONVOLUTIONAL NEURAL NETWORKS (R-CNN)

- Is an object detection method using neural networks.
- It minimizes interference by forcing the CNN to focus on a single region.
- Regions are detected by selective search algorithm.



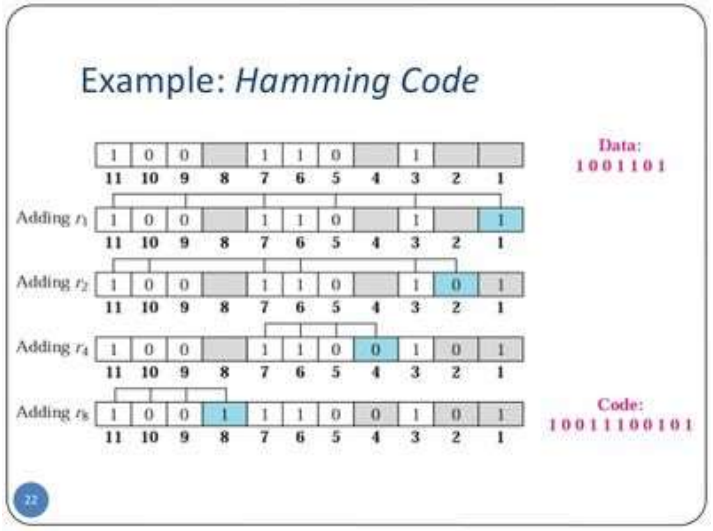
LOW-PASS

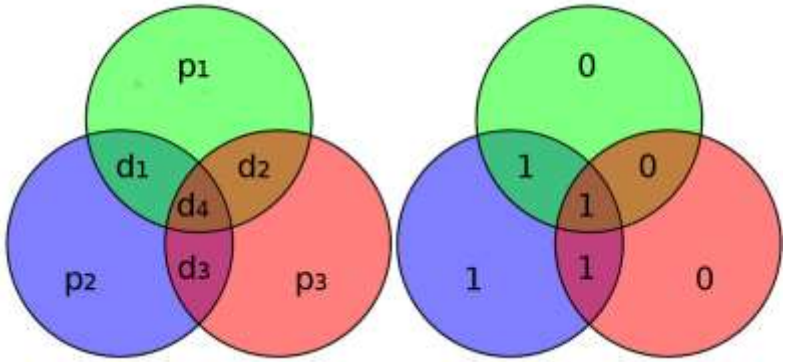
- Is a filter that passes signals with a frequency lower than a selected cutoff frequency.
- Attenuates signals with frequencies higher than the cutoff frequency.
- The exact frequency response of the filter depends on the filter design.



HAMMING

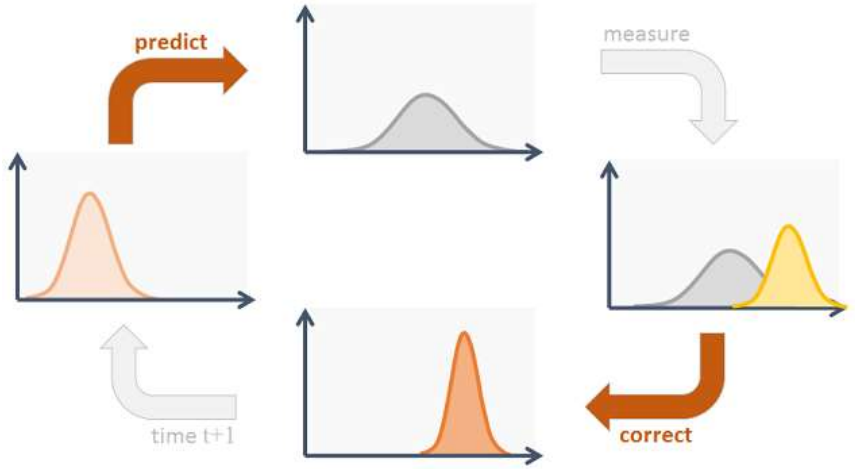
- Hamming codes are a family of linear error-correcting codes.
- Hamming codes can detect up to two-bit errors or correct one-bit errors without detection of uncorrected errors.
- They can only detect and correct errors when the error rate is low.

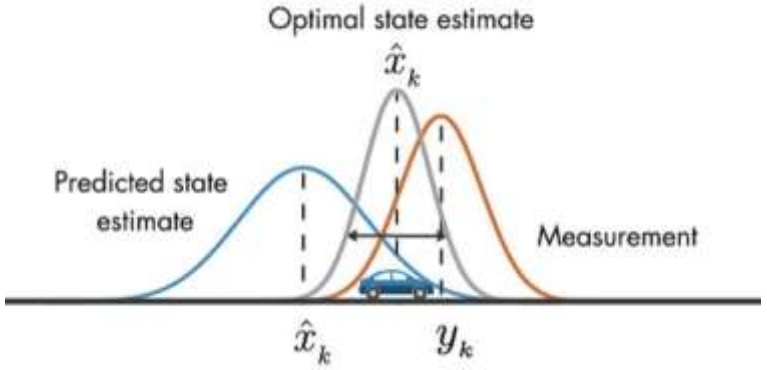




KALMAN

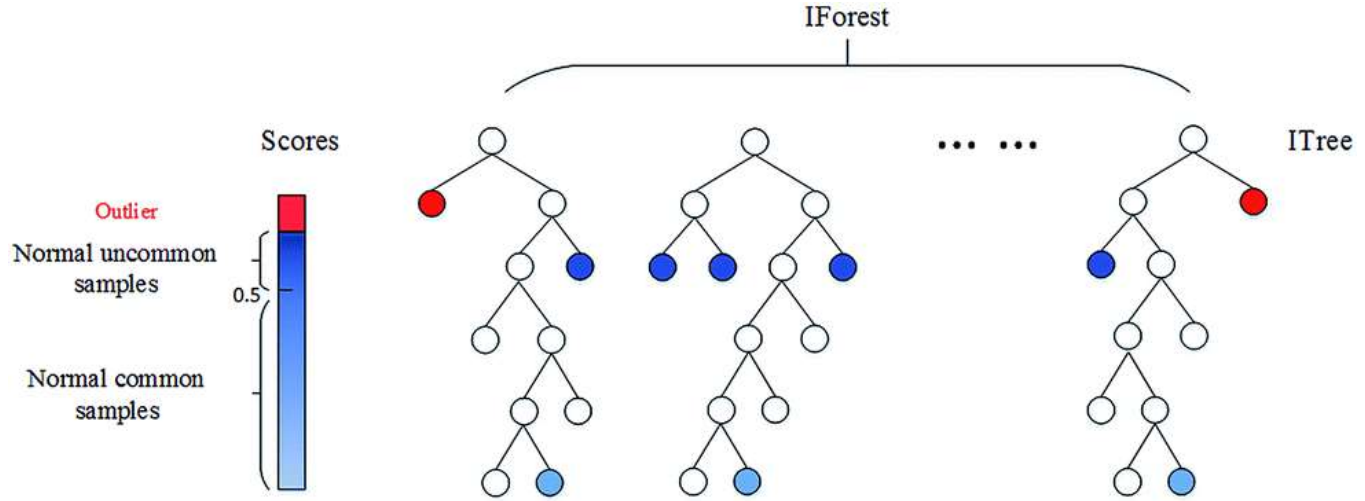
- Is an algorithm that uses a series of measurements observed over time, containing inaccuracies and produces estimates of unknown variables.
- Tend to be more accurate than those based on a single measurement.
- The Kalman filter is a widely applied concept in time series analysis and is one of the main topics in the field of robotic motion planning and control.





ISOLATION TREE

- Isolation tree is an unsupervised algorithm and therefore it does not need labels to identify the outlier/anomaly.
- The algorithm isolates each point in the data and splits them into outliers or inliers.
- An advantage of this algorithm is that it works with a huge data-set and several dimensions.



NORMALIZATION

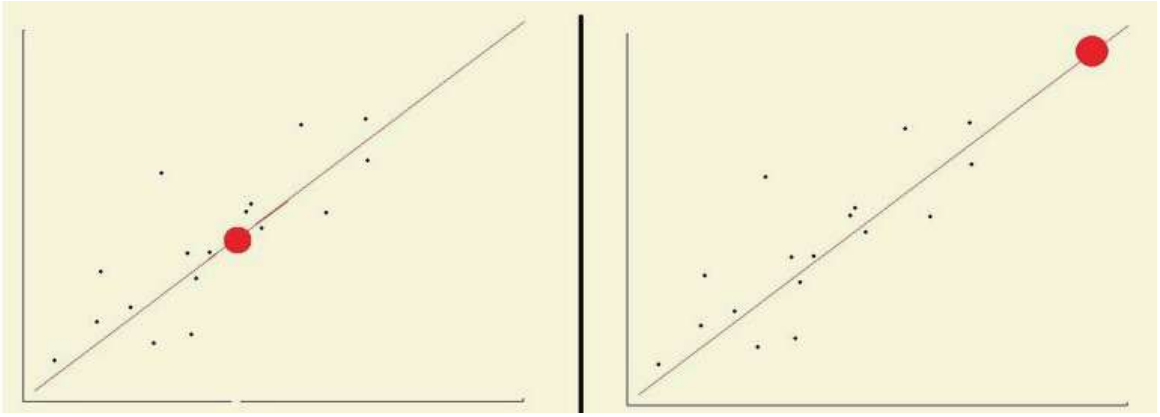
- Normalization is used to scale the data of an attribute.
- It is generally useful for classification algorithms.
- After Normalization, the data falls into a smaller range, such as -1.0 to 1.0 or 0.0 to 1.0.

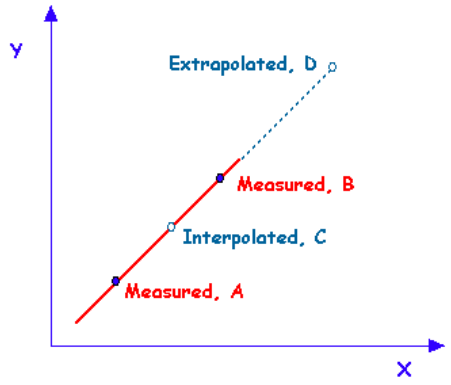
person_name	Salary	Year_of_ experience	Expected Position Level
Aman	100000	10	2
Abhinav	78000	7	4
Ashutosh	32000	5	8
Dishi	55000	6	7
Abhishek	92000	8	3
Avantika	120000	15	1
Ayushi	65750	7	5

The attributes salary and year_of_experience are on different scale and hence attribute salary can take high priority over attribute year_of_experience in the model.

INTERPOLATION & EXTRAPOLATION

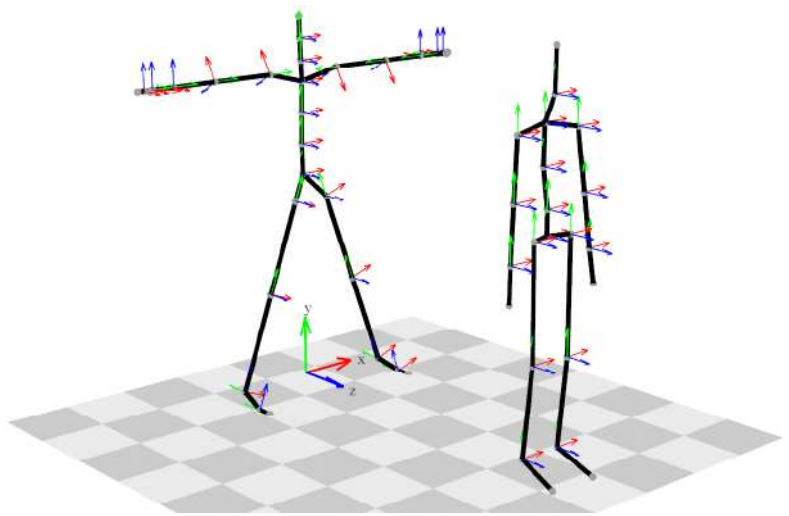
- Interpolation is a method of constructing new data points within the range of a discrete set of known data points.
- Extrapolation is an estimation of a value based on extending a known sequence of values or facts beyond the area that is certainly known.
- Both are used for predicting the value of a dependent variable.





BIOVISION HIERARCHY (BVH)

- Is a file format developed by BioVision.
- It provides skeleton hierarchy information as well as motion data.
- BVH is widely used in most 3D applications.
- It stores the animation data in terms of root, joints and offset.



R

- Is a language and environment for statistical computing and graphics.
- Provides a wide variety of statistical and graphical techniques.
- R is also an integrated suite of many software facilities (Data Manipulation, Calculation, etc…).