

# PEMBELAJARAN BAHASA ISYARAT DENGAN KINECT DAN METODE DYNAMIC TIME WARPING

**Alexander A. S. Gunawan; Ashadi Salim**

Mathematics & Statistics Department, School of Computer Science, Binus University  
Jl. K. H. Syahdan No. 9, Palmerah, Jakarta Barat 11480  
aagung@binus.edu; ashadi2048@yahoo.com

## ABSTRACT

*This study aims to develop an application program for sign language learning. The application works by detecting hand gestures using the Kinect sensor then comparing the database with the pattern of hand movements using the dynamic time warping (DTW). Kinect Sensor launched by Microsoft in 2010 is purposed for the Xbox 360 game console so that players can use body movement as game controller. The dynamic time warping (DTW) is a method for calculating the similarity between two time series data that can vary in time and speed. In tests performed by each user, there are indications that the learning curve increases along with the number of experiments conducted.*

**Keywords:** *dynamic time warping, Kinect, sign language*

## ABSTRAK

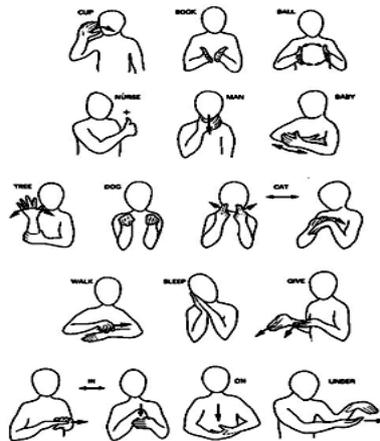
*Penelitian ini bertujuan untuk mengembangkan program aplikasi untuk pembelajaran bahasa isyarat. Cara kerja program aplikasi ini adalah dengan mendeteksi gerakan tangan dengan sensor Kinect kemudian membandingkan pola gerakan tangan tersebut dengan basis data menggunakan metode dynamic time warping (DTW). Sensor Kinect yang diluncurkan oleh Microsoft pada tahun 2010 ditujukan untuk konsol permainan XBox 360 agar pemain dapat dapat menggunakan gerakan tubuhnya sebagai pengendali permainan. Sedangkan dynamic time warping (DTW) adalah metode untuk menghitung kesamaan antara dua data deret waktu (time series) yang dapat berbeda dalam waktu dan kecepatan. Dalam pengujian yang dilakukan oleh setiap user, terdapat indikasi bahwa kurva pembelajaran meningkat seiring dengan banyaknya percobaan yang dilakukan.*

**Kata kunci:** *dynamic time warping, Kinect, bahasa isyarat*

## PENDAHULUAN

Manusia berinteraksi satu sama lain melalui komunikasi dalam bentuk bahasa. Komunikasi ini biasanya dilakukan secara verbal. Komunikasi secara verbal artinya komunikasi yang menggunakan kata-kata, baik secara lisan atau tertulis (Mulyana, D, 2005). Dalam kehidupan sehari-hari, tidak semua manusia dapat melakukan komunikasi secara lisan. Beberapa dari mereka memiliki kendala dalam hal berkomunikasi secara lisan, contohnya penderita tuna rungu dan tuna wicara. Para penderita tuna rungu dan tuna wicara memiliki cara berkomunikasi sendiri tanpa menggunakan bahasa lisan yaitu dengan bahasa isyarat.

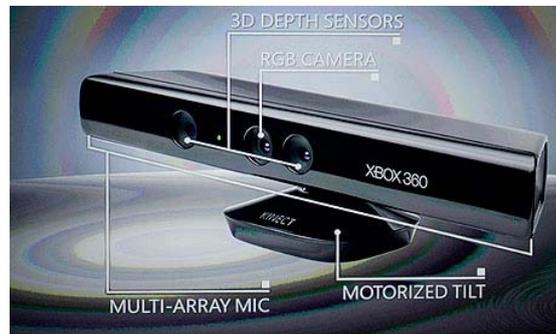
Dalam Kamus Besar Bahasa Indonesia (KBBI) (Pusat Bahasa, 2008), bahasa isyarat artinya bahasa yang tidak menggunakan bunyi ucapan manusia atau tulisan di sistem perlambangannya. Bahasa isyarat menggunakan isyarat berupa gerak jari, tangan, kepala, badan dan sebagainya, yang khusus diciptakan oleh kaum tuna rungu dan untuk kaum tuna rungu (kadang untuk kaum pendengar). Bahasa isyarat unik dalam jenisnya di setiap negara. Bahasa isyarat bisa saja berbeda di negara-negara yang berbahasa sama. Contohnya, Amerika Serikat dan Inggris meskipun memiliki bahasa tertulis yang sama, mereka memiliki bahasa isyarat berbeda. Hal yang sebaliknya juga berlaku. Ada negara-negara yang memiliki bahasa tertulis yang berbeda (contoh: Inggris dengan Spanyol), namun menggunakan bahasa isyarat yang sama. Untuk Indonesia, sistem yang sekarang umum digunakan adalah Sistem Isyarat Bahasa Indonesia (SIBI) yang sama dengan bahasa isyarat Amerika (ASL - *American Sign Language*). Pada Gambar 1 dapat dilihat sketsa contoh gerakan dasar bahasa isyarat ASL yang digunakan dalam program aplikasi yang dikembangkan.



Gambar 1 Sketsa contoh gerakan bahasa isyarat.

Bahasa isyarat tidak memiliki ruang lingkup pengguna yang besar, seperti bahasa lisan, sehingga tidak banyak orang yang dapat mengenali atau mengerti bahasa isyarat yang disampaikan lawan bicaranya. Hal ini menimbulkan kebutuhan alat bantu untuk mempelajari bahasa isyarat ini dengan mudah. Pada tahun 2010, Microsoft meluncurkan teknologi baru berupa perangkat keras sensor Kinect. Pada awalnya teknologi tersebut ditujukan sebagai konsol permainan Xbox 360 sehingga pemain dapat menggunakan gerakan tubuhnya sebagai pengendali permainan. Kinect adalah produk dari Microsoft yang awalnya dibuat khusus untuk konsol permainan Xbox 360, dimana memperkenalkan teknologi *motion gaming* sebagai fitur utamanya. *Motion gaming* maksudnya adalah membuat pemain dapat berinteraksi ketika bermain tanpa menggunakan *game controller*. Sehingga melalui Kinect, pemain dapat bermain cukup dengan menggunakan gerakan tangan atau gerakan tubuh lainnya (Metcalfe, J, 2009). Fitur yang ada pada sensor Kinect meliputi kamera RGB, *depth sensor*, *motorized tilt*, dan *multi-array microphone*. Dari keempat fitur yang dimiliki Kinect, *depth sensor*

yang berperan penting dalam sistem *motion gaming* tersebut. *Depth sensor* bertujuan untuk mendapatkan data kedalaman citra sehingga dapat dikonstruksi video tiga dimensi di dalam kondisi *ambient light* (menyesuaikan sumber cahaya yang ada di lingkungan tersebut). Pada Gambar 2 dapat dilihat bentuk dari perangkat keras Kinect. Teknologi sensor Kinect ini dapat dikembangkan lebih lanjut agar dapat digunakan untuk mengenali bahasa isyarat.



Gambar 2 Perangkat keras sensor kinect.

Dalam penelitian ini, akan dikembangkan program aplikasi untuk pembelajaran bahasa isyarat dengan menggunakan sensor Kinect untuk mendeteksi gerakan tangan dan kemudian membandingkan pola gerakan tangan tersebut dengan basis data menggunakan metode *dynamic time warping* (DTW). Sedangkan metode *dynamic time warping* (DTW) adalah metode untuk menghitung kesamaan antara dua data deret waktu (*time series*) yang dapat berbeda dalam waktu dan kecepatan (BlazStrle et al, 2003). Metode *dynamic time warping* sudah diterapkan dalam berbagai bidang, diantaranya: *handwriting matching* oleh (Tappert et al, 1990), *computer vision* dan *computer animation* oleh (Muller, M, 2007), *protein sequence alignment* dan *chemical engineering* oleh (Vial et al, 2008).

## METODE

### Sensor Kinect

Sensor Kinect dapat mendeteksi (*detection*) sendi tubuh manusia dan dilacak gerakannya (*motion tracking*). Terdapat delapan titik sendi yang digunakan yaitu: tangan kanan (*hand right*), tangan kiri (*hand left*), pergelangan tangan kanan (*wrist right*), pergelangan tangan kiri (*wrist left*), siku kanan (*elbow right*), siku kiri (*elbow left*), bahu kanan (*shoulder right*) dan bahu kiri (*shoulder left*). Pada program aplikasi yang dikembangkan ini, bahu kanan (*shoulder right*) dan bahu kiri (*shoulder left*) digunakan sebagai objek referensi yang tidak berubah selama proses pengenalan gerakan. Kedua sendi tersebut dipilih sebagai referensi karena karakteristik posisi relatifnya yang bersifat kaku (*rigid*). Selanjutnya data yang didapat perlu dinormalisasi berdasarkan titik tengah dari data kedua bahu tersebut. Hal ini dilakukan untuk menghindari perubahan nilai akibat translasi, dengan persamaan (1), dan dilasi, dengan persamaan (2), karena perubahan posisi *user* di depan kamera.

$$x_{t(\text{baru})} = x_{t(\text{lama})} - \left[ \frac{(x_{\text{bahu kiri}(\text{lama})} - x_{\text{bahu kanan}(\text{lama})})}{2} \right]$$

$$Y_{i(\text{baru})} = Y_{i(\text{jama})} - \left[ \frac{(Y_{\text{bahu kiri}}(\text{jama}) - Y_{\text{bahu kanan}}(\text{jama}))}{2} \right] \quad (1)$$

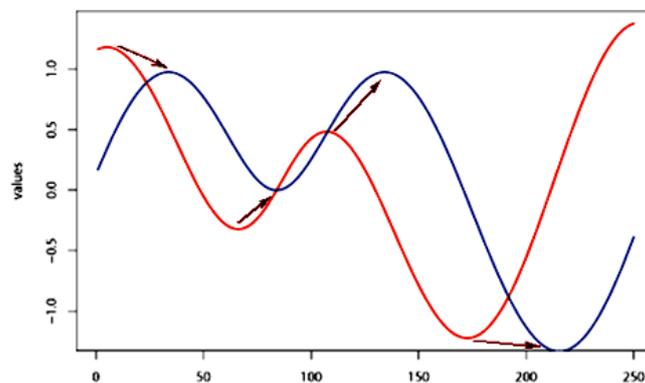
$$x_{i(\text{baru})} = \frac{x_{i(\text{jama})}}{\sqrt{(x_{\text{bahu kiri}}(\text{jama})^2 + x_{\text{bahu kanan}}(\text{jama})^2)}}$$

$$Y_{i(\text{baru})} = \frac{Y_{i(\text{jama})}}{\sqrt{(Y_{\text{bahu kiri}}(\text{jama})^2 + Y_{\text{bahu kanan}}(\text{jama})^2)}} \quad (2)$$

di mana:  $i$  = koordinat posisi tangan kanan, tangan kiri, pergelangan tangan kanan, pergelangan tangan kiri, siku kanan, dan siku kiri.

### Metode Dynamic Time Warping

Misalkan terdapat dua data deret waktu (Senin, P, 2008). yaitu  $X = (x_1, x_2, \dots, x_n), n \in \mathbb{R}$  dan  $Y = (y_1, y_2, \dots, y_m), m \in \mathbb{R}$ , dalam bentuk grafik seperti Gambar 3.



Gambar 3 Grafik perbandingan nilai X dan Y.

Pada gambar di atas, terdapat dua grafik yang menunjukkan dua data deret waktu. Dua deret waktu tersebut memiliki pola yang sama, namun letak posisi awal yang berbeda. Hal ini terjadi karena adanya perbedaan selang waktu antara  $X$  dan  $Y$ . Permasalahannya adalah bagaimana mengidentifikasi kedua data deret waktu tersebut dapat dianggap mempunyai pola yang sama walaupun memiliki dua rentang waktu dan perubahan kecepatan yang berbeda. Dengan metode *dynamic time warping*, masalah ini dapat dipecahkan dengan membandingkan biaya *optimal path warping* dengan nilai ambang batas yang ditentukan. Berikut algoritma dan kriteria yang diuraikan oleh (Senin, P, 2008) untuk menghitung *optimal path warping*:

### Menentukan Matriks Cost

Metode ini, diawali dengan menentukan matriks *cost* yang memiliki dimensi  $(n \times m)$ . Matriks *cost* adalah matriks yang merepresentasikan selisih dari himpunan  $X$  dan  $Y$ . Masing-masing elemen matriks dapat dihitung dengan persamaan berikut:

$$C_p = \begin{bmatrix} c_{1,1} & c_{1,2} & \dots & c_{1,j} \\ c_{2,1} & c_{2,2} & \dots & c_{2,j} \\ \vdots & \vdots & \ddots & \vdots \\ c_{t,1} & c_{t,2} & \dots & c_{t,j} \end{bmatrix}$$

di mana:  $C_p \in R^{n \times m}$ ;  $c_{t,j} = \|x_t - y_j\|, t \in [1:n], j \in [1:m]$

Nilai dari *cost* setiap matriksnya, digunakan untuk mencari jalur dengan *cost* terendah dari kolom pertama sampai kolom terakhir.

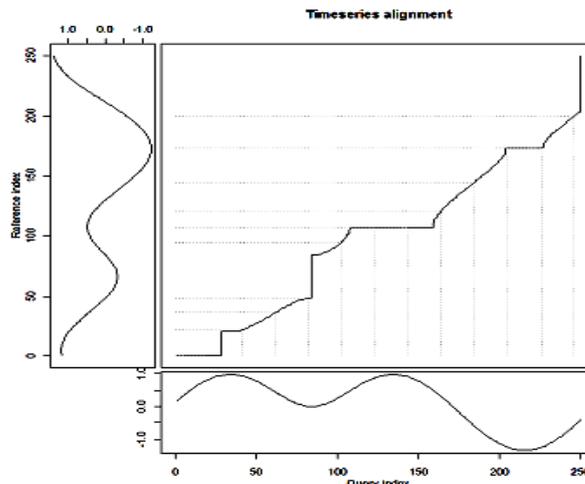
### Menentukan Optimal Warping Path

Setelah matriks *cost* terbentuk, algoritma ini akan mencari jalur *warping* dengan *cost* terendah  $p = (p_1, p_2, \dots, p_k)$  dengan  $p_i = (p_i, p_j) \in [1:n_i] \times [1:m_i]$  dan  $i \in [1:k]$ . *Optimal warping path* seperti pada Gambar 4 adalah jalur dengan biaya (*cost*) terpendek dari selisih perbandingan X dan Y. Dengan jalur *warping* yang akan dicari harus memenuhi kriteria-kriteria sebagai berikut:

*Boundary condition*:  $p = (1,1)$  dan  $p = (n,m)$ . Awal dan akhir jalur harus merupakan titik awal dan titik akhir dari deret waktu.

*Monotonicity condition*:  $n_1 \leq n_2 \leq \dots \leq n_k$  dan  $m_1 \leq m_2 \leq \dots \leq m_k$ . kondisi ini akan mempertahankan deret urutan waktu.

*Step size condition*: membatasi pergeseraan *path* atau jalur yang terlalu jauh.



Gambar 4 Optimal path warping dari dua deret waktu X dan Y.

### Menentukan Matriks DTW

Berdasarkan definisi dari *optimal warping path* dalam pencarian jalur dengan matriks *cost*, diperlukan pengujian ke setiap jalur antara X dan Y. Hal ini tentu saja tidak efisien, karena kemungkinan jalur yang berpotensi untuk dilewati juga akan semakin banyak seiring dengan panjang deret waktu X dan Y. Untuk mengatasi masalah ini, digunakan *dynamic programming* sebagai metode optimasi pencarian jalur dengan cepat. Menurut (Levitin, A, 2006) *dynamic programming* adalah sebuah teknik untuk menyelesaikan masalah dengan cara membagi masalah dalam beberapa sub masalah (*stage*) sehingga solusi dari masalah dapat dipandang sebagai serangkaian keputusan yang saling berkaitan. Dalam kasus ini dilakukan perhitungan untuk total *cost* setiap pencarian jalur yang

mungkin dilewati (*stage*). Semua jalur yang telah dilewati diakumulasi terlebih dahulu. Matriks akumulasi yang akan digunakan didefinisikan sebagai berikut:

$$DTW(X, Y) = C_p(X, Y)$$

di mana perhitungan akumulasi untuk matriks DTW adalah sebagai berikut:

Baris pertama:  $DTW(1, j) = \sum_{k=1}^j c(x_1, y_k), j \in [1, m]$

Kolom pertama:  $DTW(i, 1) = \sum_{k=1}^i c(x_k, y_1), i \in [1, n]$

Semua elemen matriks:

$$DTW(i, j) = \min\{DTW(i-1, j-1), DTW(i-1, j), DTW(i, j-1)\} + c(x_i, y_j), i \in [1, n], j \in [1, m]$$

### Menentukan Cost Optimal Warping Path

Setelah menghitung matriks DTW, untuk menentukan total dari cost jarak yang melewati jalur *warping* dapat menggunakan persamaan berikut:

$$cost = \frac{1}{m} \{\min[DTW(1, m), DTW(2, m), \dots, DTW(n, m)]\}$$

dengan *cost* dari *optimal warping path* adalah *cost* akumulasi minimum pada kolom terakhir.

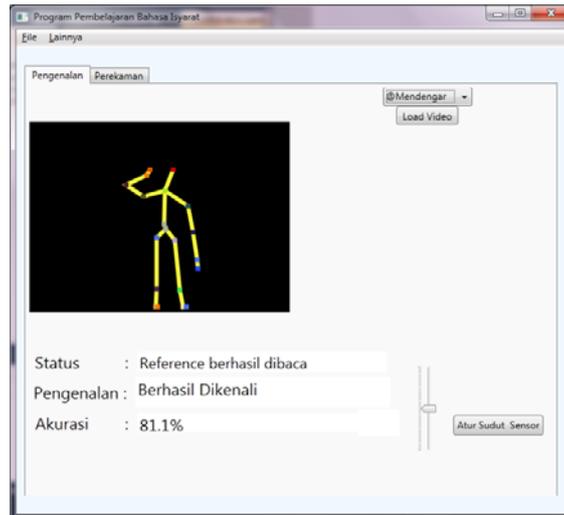
### Pengembangan Aplikasi

Program aplikasi yang dikembangkan terdiri dari dua menu utama, yaitu: Pengenalan dan Perekaman. Menu pengenalan bertujuan untuk menerjemahkan bahasa isyarat yang dilakukan *user* di depan sensor Kinect menjadi hasil penerjemahan dalam bentuk tulisan. Untuk melakukan proses pengenalan pola gerakan bahasa isyarat, dilakukan langkah-langkah sebagai berikut: (1) pengambilan gerakan dengan sensor Kinect; (2) pengenalan objek kerangka; (3) normalisasi data; (4) pengenalan dengan metode *dynamic time warping*.

Sedangkan menu perekaman bertujuan untuk merekam pola gerakan yang dilakukan user di depan sensor Kinect dalam bentuk *file*, dan mengubah pola gerakan tersebut menjadi basis data yang berupa nilai vektor koordinat titik sendi. Proses perekaman itu sendiri meliputi langkah-langkah sebagai berikut: (1) pengambilan gambar dengan sensor Kinect; (2) pengenalan objek kerangka; (3) normalisasi data; (4) *generate data reference*.

## HASIL DAN PEMBAHASAN

Gambar 5 berikut ini merupakan *screenshot* dari program aplikasi yang dikembangkan:



Gambar 5 Screenshoot program aplikasi.

Untuk pengujian program aplikasi yang dikembangkan, direkam enam pola gerakan bahasa isyarat sebagai data referensi dan kemudian gerakan tersebut ditiru oleh lima *user* dalam tiga kali percobaan. Dengan menetapkan nilai ambang batas 0.6, didapatkan dari hasil tiga kali percobaan meniru pola gerakan sebagai berikut (Tabel 1):

Tabel 1 Tabel persentase pengenalan pola gerakan

No	Pola gerakan	Persentase pengenalan
1	Saya	100%
2	Bisa	100%
3	Mendengar	100%
4	Melihat	100%
5	Dengan	60%
6	Membaca	46.6%

Enam pola gerakan bahasa isyarat ini mempunyai tingkat kesulitan yang berbeda-beda. Sehingga dapat dilihat pada pola gerakan yang rumit yaitu “Dengan” dan “Membaca”, *user* kesulitan dalam meniru gerakan bahasa isyaratnya. Gerakan “Dengan” agak sulit ditirukan karena orientasi posisi tangan tidak biasa dilakukan oleh *user*. Sedangkan gerakan “Membaca” sulit ditirukan karena membutuhkan gerakan jari tangan. Pada gerakan “Membaca” ini, program aplikasi yang dikembangkan belum mampu melacak gerakan jari tangan, sehingga kemungkinan bahwa kesalahan pengenalan ada pada program aplikasi memang besar. Untuk pola gerakan yang mudah yaitu “Saya”, “Bisa”, “Mendengar” dan “Melihat”, *user* dapat meniru gerakan dengan mudah. Dalam 3 kali percobaan yang dilakukan oleh setiap *user*, terdapat indikasi bahwa kurva pembelajaran (*learning curve*) meningkat seiring dengan banyaknya percobaan yang dilakukan.

## SIMPULAN

Kombinasi antara perangkat keras sensor Kinect dan metode *dynamic time warping* yang digunakan dalam penelitian ini mampu untuk mengidentifikasi pola gerakan berbasis kerangka tubuh manusia dan dapat diterapkan untuk mengenali bahasa isyarat. Selanjutnya pengujian, apakah program

aplikasi ini dapat membantu pembelajaran bahasa isyarat mengindikasikan kurva pembelajaran (*learning curve*) meningkat seiring dengan banyaknya percobaan yang dilakukan dan tergantung pula pada tingkat kesulitan gerakan bahasa isyarat tersebut bagi *user*. Untuk pengembangan lebih lanjut, program aplikasi pembelajaran bahasa isyarat ini akan mampu mendeteksi jari tangan (*finger*) dan melacak gerakan jari tangan (*finger*) tersebut karena banyak kosakata dalam bahasa isyarat dimana gerakan jari tangan berperan penting.

## DAFTAR PUSTAKA

- BlazStrle, Mozina, M., Bratko, I. (2003). *Qualitative Approximation to Dynamic Time Warping Similarity between Time Series Data*. University of Ljubljana. Slovenia.
- Levitin, A. (2006). *Introduction to the Design & Analysis of Algorithms* (2nd ed.). Boston: Addison Wesley.
- Metcalf, J. (2009, Juni). *E3 2009: Microsoft at E3 Several Metric Tons of Press Releaseapalloza*. Retrieved Desember 31, 2011 From Digital Joystick: Video Game News & Views <http://blog.seattlepi.com/digitaljoystick/2009/06/01/e3-2009-microsoft-at-e3-several-metric-tons-of-press-releaseapalloza/>.
- Muller, M. (2007). DTW-based motion comparison and retrieval. *Information Retrieval for Music and Motion Part II*, 211-226. New York City: Springer.
- Mulyana, D. (2005). *Ilmu Komunikasi: Suatu Pengantar*. Bandung: Remaja Rosdakarya.
- Pusat Bahasa. (2008). *Kamus Besar Bahasa Indonesia*. Jakarta: Gramedia Pustaka Utama.
- Senin, P. (2008). *Dynamic Time Warping Algorithm Review*. University of Hawaii at Manoa, Honolulu, USA.
- Tappert, C. C., Suen, C. Y., Wakahara, T. (1990) The state of the art on online handwriting recognition. *IEEE Transactions* 12 (8), 787.
- Vial, J., Nocairi, H., Sassi, P., Mallipatu, S., Cognon G., Thiebaut, D., Teillet, B. dan Rutledge, D. (2008). Combination of dynamic time warping and multivariate analysis for the comparison of comprehensive two-dimensional gas chromatograms application to plant extracts. *Journal of Chromatography*.