BAB II

METODE DESKRIPTOR BENTUK DARI CITRA DENTAL X-RAY

Pada bab ini akan dijelaskan mengenai metode-metode yang digunakan dalam membentuk deskriptor bentuk dari citra *dental x-ray* dan mengukur derjat kemiripan antar citra *dental x-ray*.

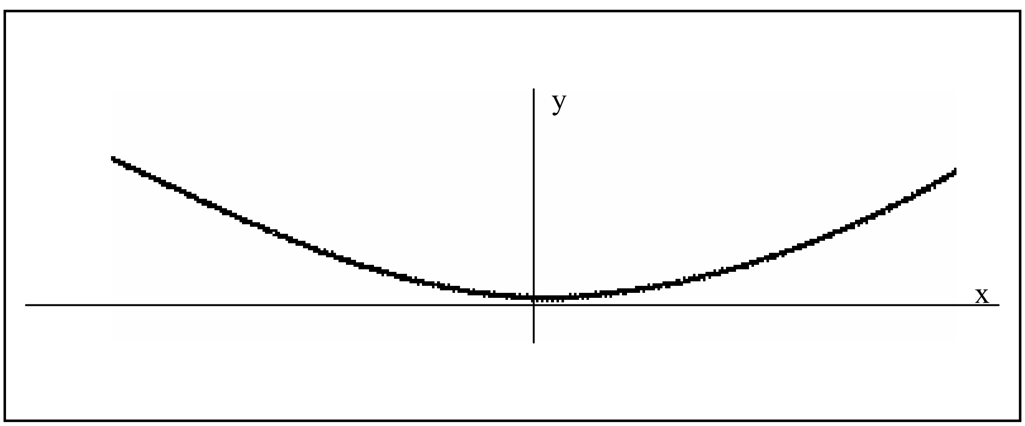
2.1 DESKRIPTOR BENTUK

Secara umum, deskriptor bentuk merupakan himpunan dari nilai-nilai yang dihasilkan untuk mendeskripsikan sebuah bentuk. Nilai dari sebuah deskriptor bergantung pada bentuk yang dideskripsikan. Makin mampu sebuah deskriptor dalam membedakan sebuah bentuk dari bentuk lainnya, maka makin bagus deskriptor tersebut [3]. Namun pada kenyataannya sebuah deskriptor tidak selalu mampu merekosntruski sebuah bentuk yang sama persis dengan bentuk awalnya, dengan kata lain sebuah deskriptor hanya melakukan aproksimasi dalam mendeskripsikan sebuah bentuk.

Secara garis besar, deskriptor bentuk bisa dibedakan menjadi dua jenis yaitu *boundary descriptor* dan *region descriptor* [3]. *Boundary descriptor* dalam melakukan prosesnya tidak melibatkan keseluruhan piksel, tetapi hanya melibatkan piksel-piksel yang berada di tepi-tepi sebuah objek yang akan dideskripsikan. Sebaliknya, *region descriptor* melibatkan seluruh piksel pada sebuah objek dalam melakukan penghitungan deskriptor.

Pemilihan dari deskriptor bentuk sangat bergantung kepada hasil dari proses ekstraksi ciri. Pada penelitan ini informasi ciri yang digunakan untuk melakukan proses pencocokan adalah bentuk rahang bawah dan bentuk dagu. Kedua ciri bentuk tersebut merupakan ciri yang bersifat *boundary-based*, maka dalam pembentukan deskriptornya metode-metode yang digunakan adalah metode-metode yang mampu mendeskripsikan bentuk berdasarkan informasi-informasi yang bersifat *boundary-based*.

Deskriptor bentuk yang digunakan pada penelitian ini ada dua jenis, yaitu metode *quadratic regession* dan *centroid distance.* Untuk mendeskripsikan hasil ekstraksi yang berbentuk lengkung dagu pendekatan yang dapat digunakan adalah dengan menggunakan metode *quadratic regession*. Alasannya, bentuk lengkung dagu dapat dianggap sebagai sekumpulan data yang membentuk sebuah fungsi yang kuadratik yang nilainya relatif terhadap sebuah sumbu imajiner x dan y, sebagaimana yang ditunjukkan pada gambar 2.1.

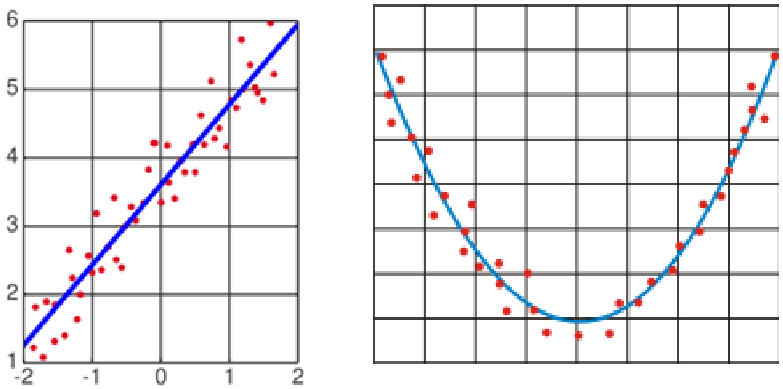


Gambar 2.1 Hasil ekstraksi bentuk lengkung dagu

Sedangkan untuk hasil ekstraksi yang berbentuk lengkung rahang, pendekatan deskriptor yang digunakan adalah dengan menggunakan metode *centroid distance*. Metode *centroid distance* dapat digunakan sebagai deskriptor bentuk dari citra *dental x-ray* karena bentuk dari lengkung rahang dapat dianggap sebagai sebuah kurva yang tertutup, sehingga memungkinkan untuk menemukan nilai dari jarak titik pusat.

2.1.1 Deskriptor Bentuk dengan Pendekatan *Quadratic Regression*

Metode regresi (*regression*) pada dasarnya merupakan pendekatan statistik yang digunakan untuk melakukan estimasi dan analisis data dengan cara melakukan generalisasi terhadap data yang terbatas. Proses generalisasi dilakukan dengan cara memetakan data-data yang ada ke dalam sebuah fungsi. Bentuk fungsi yang digunakan bergantung pada bentuk sebaran data. Untuk data yang tersebar secara linear, maka fungsi yang digunakan adalah fungsi linear (lihat gambar 2.2 a), sedangkan untuk data yang sebaranya mendekati nilai yang kuadratik maka fungsi yang di gunakan adalah fungsi kuadratik pula (lihat gambar 2.2 b) [4].



(a) (b)

Gambar 2.2 Sebaran data(a) pada *linear regression* (linear regression), (b) pada *quadratic regression*

Persamaan dari linear regression ditunjukkan pada persamaan berikut:

$$y_{(x,w)} = w_0 + w_1 x_1 + \ldots + w_n x_n \qquad (2.1)$$

Dimana x = (x_1 ,x_2 ,…,x_n) merupakan sebuah vektor yang berisi data-data variabel input. Sedangkan w_0 ,......, w_n parameter dari fungsi linear di atas. Fungsi dari persamaan 2.1 bisa diperluas dengan cara menjadikan varibel input x_i sebagai nilai dari sebuah fungsi basis (*basis functions*) [4]. Berdasarkan hal tersebut persamaan 2.1 bisa di tulis sebagai berikut:

$$y(x,w) = w_0 + \sum_{j=1}^{M-1} w_j \phi_j(x) \qquad (2.2)$$

Dimana $\phi_j(x)$ merupakan sebuah fungsi basis (*basis function*). Dengan menunjukkan nilai maksimum dari indeks j adalah M-1 maka jumlah parameter pada persamaan tersebut adalah M. Parameter w_0 pada persamaan 2.2 sering disebut sebagai nilai bias. Persamaan 2.2 bisa lebih disederhanakan dengan menambahkan sebuah fungsi basis tambahan $\phi_0(x) = 1$ [4] sehingga persamaan 2.2 menjadi

$$y(\mathrm{x},\mathrm{w}) = \sum_{j=0}^{M-1} w_j \phi_j\ (\mathrm{x}) = \mathrm{w}^{\mathrm{T}} \phi(\mathrm{x}) \tag{2.3}$$

dimana w = $(w_0, \ldots, w_{M-1})^{\mathrm{T}}$ dan $\phi = (\phi_{0, \ldots,} \phi_{M-1})^{\mathrm{T}}$.

Fungsi pada persamaan 2.2 dikatakan sebagai linear model dikarenakan fungsi tersebut linear terhadap w. Dengan kata lain sebuah fungsi disebut sebagai linear regresi jika derajat tertinggi dari fungsi basis nya bernilai 1.

Fungsi *linear regression* bisa dikembangkan menjadi sebuah fungsi regresi yang polynomial dengan cara menjadikan fungsi basis sebagai sebuah fungsi pangkat, dimana $\phi_j(\mathrm{x}) = x^j$.

Berdasarkan definisi tersebut, maka sebuah fungsi *quadratic regression* yang merupakan bagian dari fungsi *polynomial regression* bisa didefinisikan sebagai kombinasi dari parameter w_j terhadap fungsi basis $\phi_j(\mathrm{x}) = x^j$ dimana nilai j tertinggi adalah 2 [4], seperti yang di tunjukkan oleh persamaan berikut :

$$y(\mathrm{x},\mathrm{w}) = w_0 + w_1 \phi_1(\mathrm{x}) + w_2 \phi_2(\mathrm{x}) = \sum_{j=0}^{2} w_j \phi_j(\mathrm{x}) = \mathrm{w}^{\mathrm{T}} \phi(\mathrm{x}) \tag{2.4}$$

Untuk data input yang jumlahnya banyak, persamaan 2.4 bisa diperluas menjadi :

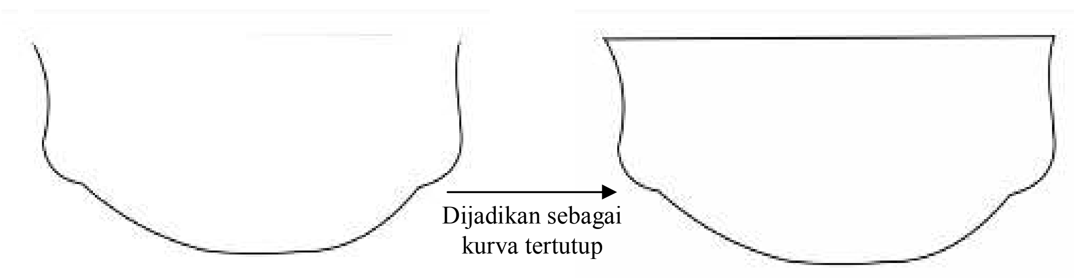
$$\begin{bmatrix} y_1 \\ : \\ y_n \end{bmatrix} = \begin{bmatrix} w_0 + w_1 x_1 + w_2 {x_1}^2 \\ : \\ w_0 + w_1 x_n + w_2 {x_n}^2 \end{bmatrix} \tag{2.5}$$

$$\begin{bmatrix} y_1 \\ : \\ y_n \end{bmatrix} = \begin{bmatrix} 1 & x_1 & {x_1}^2 \\ : & : & : \\ 1 & x_n & {x_n}^2 \end{bmatrix} \begin{bmatrix} w_0 \\ w_1 \\ w_2 \end{bmatrix} \tag{2.6}$$

$$\mathrm{y} = \mathbf{XW} \tag{2.7}$$

2.2.1 Deskriptor Bentuk dengan Pendekatan C*entroid Distance*

Deskriptor bentuk dengan *centroid distance* (jarak titik pusat) digunakan untuk mendeskripsikan hasil dari ekstraksi ciri *dental x-ray* yang berupa bentuk lengkung rahang secara keseluruhan. Bentuk lengkung ini pada dasarnya merupakan sebuah kurva yang terbuka, tetapi agar bisa diaplikasikan terhadap deskriptor jarak titik pusat maka kurva terbuka ini bisa dianggap sebagai sebuah kurva tertutup, seperti yang ditunjukkan oleh gambar 2.3.



Gambar 2. 3 Perubahan kurva terbuka lengkung rahang menjadi kurva tertutup

Pada metode jarak titik pusat yang didefinisikan sebagai jarak titik pusat adalah jarak dari pinggir kurva ke titik pusat kurva. Nilai titik pusat dari sebuah kurva tertutup dapat dicari dengan mengitung titik berat dari kurva tersebut, seperti yang ditunjukkan oleh persamaan 2.5.

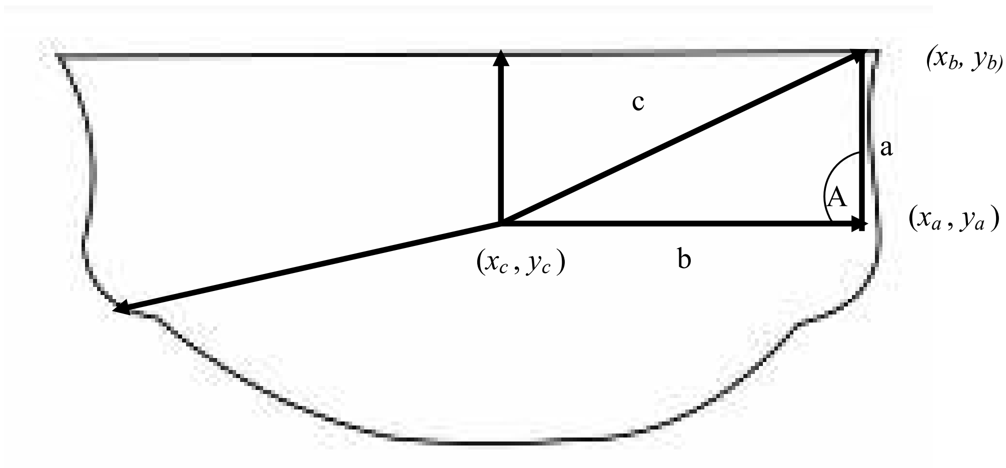
$$[x_c\ y_c] = \sum_{i=0}^{n} [x_i\ y_i]/n \qquad (2.5)$$

$[x_c\ y_c]$ merupakan sebuah vektor yang merepresentasikan posisi dari titik pusat sebuah kurva. Sedangkan *n* merupakan jumlah dari seluruh piksel yang membentuk kurva tersebut.

Jarak titik pusat ditunjukkan oleh persamaan berikut:

$$r(t) = \left(\left[x(t) - x_c\right]^2 + \left[y(t) - y_c\right]^2\right)^{1/2} \qquad (2.6)$$

r(t) merupakan jarak titik pusat terhadap titik t yang berada pada kurva, sedangkan x_c dan y_c merupakan titik pusat dari kurva yang nilainya didapat lewat persamaan 2.6.



Gambar 2. 4 Cara merepresentasikan bentuk rahang dengan metode *centroid distance*

Langkah awal, adalah menentukan titik awal (x_a , y_a) pada kurva. Agar dalam perhitungannya konsisten, maka posisi titik awal untuk setiap deskriptor diorientasikan searah 180^0 terhadap titik pusat. Dalam penghitungan jarak titik pusat, agar mengurangi komputasi maka tidak semua titik pada kurva yang diukur, tetapi hanya mengambil beberapa titik dengan teknik *sampling* berdasarkan sudut. Besarnya sudut diperoleh dengan menggunanakan aturan *cos*, seperti yang ditunjukkan oleh persamaan dibawah ini:

$$c^2 = a^2 + b^2 - ab \cos (A) \qquad (2.7)$$

BAB III

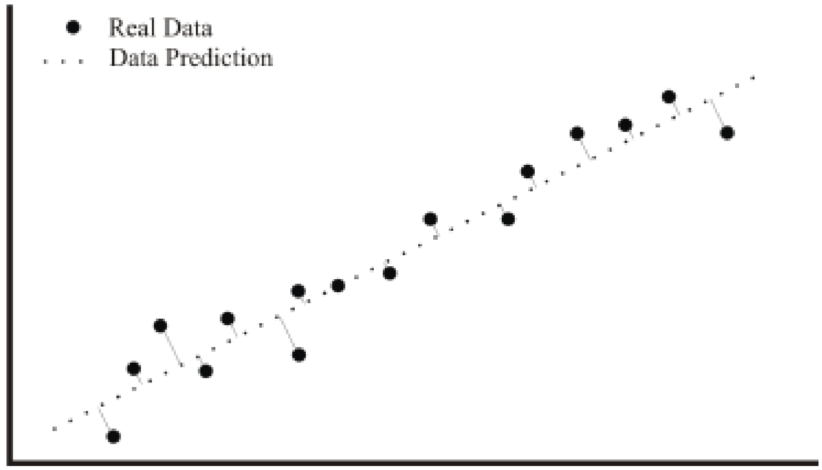
METODE PENGUKUR KEMIRIPAN

Pada bagian ini akan dijelaskan metode yang digunakan untuk mengukur kemiripan dua buah citra *x-ray* rahang. Proses pengukuran kemiripan merupakan kelanjutan dari proses pembentukan deskriptor yang telah dibahas pada bab II.

Metode yang digunakan pada proses pengukuran kemiripan, sangat bergantung pada jenis deskriptor yang digunakan. Untuk deskriptor yang menggunakan metode jarak titik pusat metode pengukuran yang digunakan adalah metode *fuzzy similarity.* Sedangkan untuk deskriptor yang menggunakan pendekatan *quadratic regression*, metode yang digunakan adalah dengan menggunakan metode pengukuran dengan *least square loss function.*

3.1 PENGUKUR KEMIRIPAN DENGAN *LEAST SQUARE LOSS FUNCTION*

Hasil data yang dimodelkan dengan menggunakan metode *regression,* baik itu *linear regression* ataupun *quadratic regression* tidak selalu sama dengan himpunan data yang sebenarnya, seperti yang diilustrasikan pada gambar 3.1. Di ilustrasi tersebut, lingkaran hitam merupakan data yang sebenarnya dan titik-titik yang ada pada garis merupakan data yang diprediksi lewat sebuah fungsi regresi.



Gambar 3. 2 Ilustrasi dari error pada data prediksi

Selisih antara data yang sebenarnya dan data hasil prediksi dikenal dengan istilah *residuals. Residuals* dari sebuah model regresi ditunjukkan oleh persamaan berikut [4][5]:

$$r_i = y_i - \sum_{j=1}^{M-1} w_j \phi_j \ (\mathrm{x}) \ , i = 1,...,m \tag{3.1}$$

Dimana r_i merupakan nilai residual, y_i merupakan data ke-i yang sebenarnya, dan $\sum_{j=1}^{M-1} w_j \phi_j \ (\mathrm{x})$ merupakan fungsi dari data prediksi dengan $\phi(\mathrm{x})$ merupakan sebuah fungsi basis. Persamaan 3.1 jika diubah dalam bentuk notasi vektor matrik maka akan berbentuk:

$$\begin{bmatrix} r_1 \\ : \\ r_n \end{bmatrix} = \begin{bmatrix} y_1 \\ : \\ y_n \end{bmatrix} - \begin{bmatrix} 1 & x_1 & x_1^{\ 2} \\ : & : & : \\ 1 & x_n & x_n^{\ 2} \end{bmatrix} \begin{bmatrix} w_0 \\ w_1 \\ w_2 \end{bmatrix} \tag{3.2}$$

Keakuratan pemodelan data lewat metode regressi berbanding lurus dengan minimalisasi dari nilai *residual,* dengan kata lain semakin kecil nilai residual dari sebuah model berarti semakin akurat model tersebut. Cara untuk menghitung nilai dari residual adalah dengan cara *least square* [6]:

$$\|r\|^2 = \sum_{1}^{m} r_i^{\,2} \tag{3.3}$$

Perlu dicari nilai parameter yang sesuai, sehingga bisa ditemukan nilai residual yang minimal. Berdasarkan persamaan 2.7, maka nilai dari parameter dapat dicari dengan persamaan :

$$\mathbf{W} = \mathbf{X}^{-1}\, \mathrm{y} \tag{3.4}$$

Berdasarkan persamaan 2.6 diketahui bahwa $\mathbf{X} = \begin{bmatrix} 1 & x_1 & x_1^{\,2} \\ : & : & : \\ 1 & x_n & x_n^{\,2} \end{bmatrix}$ oleh karena itu persamaan 3.4 hanya bisa diselesaikan jika matrik **X** merupakan sebuah matriks bujur sangkar, karena selain dari matriks bujur sangkar tidak bisa dicari matriks *inverse*-nya. Untuk matriks **X** yang tidak bujur sangkar maka persamaan 3.4 menjadi [7]:

$$\mathbf{W} = (\mathbf{X}^{\mathrm{T}}\, \mathbf{X})^{-1}\, \mathrm{y} \tag{3.5}$$

3.2 PENGUKUR KEMIRIPAN MENGGUNAKAN *FUZZY SIMILARITY*

Pengukuran dengan metode ini digunakan dalam menjadi derajat kemiripan sebuah citra berdasarkan deskriptor *centroid distance*. Pada metode deskriptor *centroid distance* seluruh jarak titik pusat dari sebuah citra rahang disimpan kedalam sebuah array. Data-data jarak titik pusat inilah yang nantinya dijadikan sebagai pola masukan dan pola prototipe *fuzzy*. Pola masukan *fuzzy* diambil dari jarak titik pusat citra input, sedangkan pola prototipe diambil dari jarak titik pusat seluruh citra yang ada di basis data.

Proses pengukuran kemiripan dengan menggunakan metode *fuzzy,* sudah sering digunakan dalam bidang *content based image retrieval.* Pengukuran kemiripan *fuzzy* dilakukan dengan cara mengukur kemiripan antara pola masukan dan pola prototipe *fuzzy.* Hasil dari proses pengukuran kemiripan disimpan ke dalam sebuah vektor [8]:

$$S(q,t_i) = [S(q,t_{i1}), S(q,t_{i2}), \ldots, S(q,t_{in})]^T \tag{3.6}$$

Dimana q merupakan pola masukan *fuzzy,* t_i merupakan pola prototipe *fuzzy*, $S(q_i, t_{ij})$ merupakan derajat kemiripan dari elemen ke-*j,* dan n merupakan dimensi dari pola masukan *fuzzy* (q). Hasil dari $S(q_i, t_{ij})$ bukanlah sebuah skalar, untuk mendapatkan nilai skalar dari $S(q_i, t_{ij})$ maka perlu ditambahkan sebuah bobot w_j dimana nilainya tidak selalu sama untuk setiap elemen, seperti yang ditunjukkan oleh persamaan berikut:

$$S(q, \quad t_i) = \sum_{j=1}^{n} w_j \mathrm{S}(\mathrm{q_i}, t_{ij}) \tag{3.7}$$

Selanjutnya, nilai dari derajat kemiripan setiap elemen harus dipetakan kedalam sebuah *fuzzy membership function* ($\mu_A(x)$) yang nilainya merupakan bilangan real positif antara [0,1].

$$0 \leq \quad \mu_A(x) \ \leq \ 1 \tag{3.8}$$

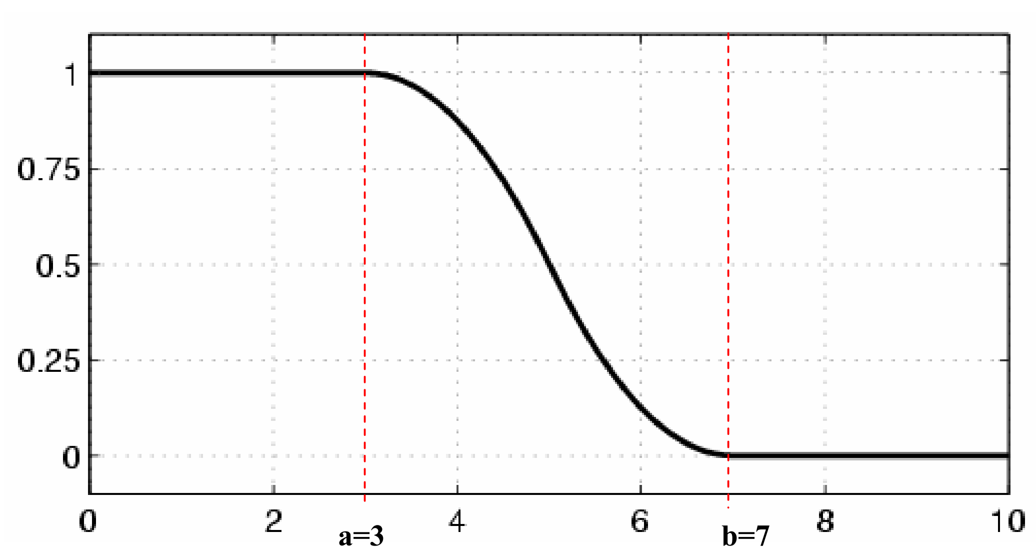
Definisi formal dari sebuah *Fuzzy membership function* ditunjukkan oleh persamaan berikut:

$$\mu_A : x \to [0,1], x \in X \tag{3.9}$$

Dimana x merupakan bilangan real yang mendeskripsikan hasil pemetaan dari sebuah objek, X adalah semesta pembicaraan dan A adalah subset dari X [9].

Jenis *membership function* yang sering digunakan beragam sekali, tergantung dari jenis objek yang akan dipetakan. Di antara jenis membership function yang sering digunakan seperti, *triangular*, *Gaussian*, *S-function*, dan zmf [9].

Pada penelitian ini, jenis membership function yang digunakan adalah jenis zmf. *Membership function* jenis ini merupakan fungsi berbasis spline dari *x*. Disebut sebagai zmf karena kurva yang dibentuk menyerupai huruf z (lihat gambar 3.2) .



Gambar 3. 3 Fungsi zmf dengan [a,b] = [3,7]

Fungsi keanggotaan dari zmf ditunjukkan oleh persamaan berikut, variabel x merupakan informasi dari deskriptor yang akan diukur kemiripannya:

$$y = \left\{ \begin{array}{c} 1 \text{ , } x \le a \\ 1-2\left(\dfrac{x-a}{b-a}\right)^2 \text{, } a \le x \le \dfrac{a+b}{2} \\ 2\left(b-\dfrac{x}{b-a}\right)^2 \text{, } \dfrac{a+b}{2} \le x \le b \\ 0 \text{ , } \quad x \ge b \end{array} \right\} \qquad (3.10)$$