

BAB II

TINJAUAN PUSTAKA

Pada bab dua ini, akan dibahas mengenai transformasi wavelet dan penurunannya serta proses dalam melakukan Fuzzy Subspace Cluster.

2.1. Sensor Vibrasi Piezoelektrik

Perubahan variasi sifat material dapat ditinjau sebagai gambaran umum dari hubungan antara sifat kimia dengan sensor. Masa dan kecepatan adalah sifat yang penting untuk sensor piezoelektrik. Oleh karena itu microbalances dan microviscometers, menggunakan kristal piezoelektrik, dinyatakan sebagai suatu sistem sensor. Kompresi dari suatu kristal quartz menghasilkan suatu potensial listrik. Satu prinsip yang dapat meninjau efek ini untuk membangkitkan gelombang akustik pada bendap padat dengan mengaplikasikan potensial listrik bolak-balik ke suatu material piezoelektrik ditunjukkan pada gambar 1. Gelombang akustik, khususnya frekuensi dan resonant resistance, sangat dipengaruhi oleh kondisi batas yang dibentuk oleh dimensi fisik dari alat dan oleh sifat fisik dari material disepanjang lintasan gelombang.

Sensor piezoelektrik adalah peralatan elektronik pasif berfase padat (solid-state) yang dapat merespon perubahan temperature, tekanan, dan yang paling penting merespon sifat fisik (physical properties) pada suatu interface antara permukaan alat dan fluida atau padatan asing. Perubahan pada sifat fisik antara lain seperti masa jenis, kelistrikan, viskositas, dan ketebalan lapisan. Sensor piezoelektrik beroperasi dengan mengobservasi penyebaran dari suatu gelombang akustik melalui solid-state device. Deteksi sensor dilakukan dengan meninjau korelasi variasi penyebaran gelombang akustik ke sejumlah perekam analyte pada

permukaan dan kemudian ke konsentrasi analyte di dalam sampel yang tertangkap sensor atau dikorelasikan dengan perubahan pada sifat fisik dari interfacial thin films.



Gambar 2.1. Sensor Vibrasi Piezoelektrik

Pada gambar 2.1. diatas sensor dapat bekerja pada getaran sampai dengan 40 Hz atau 2.400 rpm dengan high voltage sensivity $1 \text{ V/g} - 5 \text{ V/g}$ pada arah vertical maupun horizontal. Output sensor sangat kecil sehingga perlu adanya penguatan terlebih dahulu sebelum di konversi kedalam ADC (analog to digital converter)

2.2. Penguat Op-Amp

merupakan suatu jenis penguat elektronika dengan sambatan arus searah yang memiliki bati (faktor penguatan atau dalam bahasa Inggris: gain) sangat besar dengan dua masukan dan satu keluaran. Penguat operasional pada umumnya tersedia dalam bentuk sirkuit terpadu dan yang paling banyak digunakan adalah seri 741.

Penguat operasional adalah perangkat yang sangat efisien dan serba guna.[3] Contoh penggunaan penguat operasional adalah untuk operasi matematika sederhana seperti penjumlahan dan pengurangan terhadap tegangan

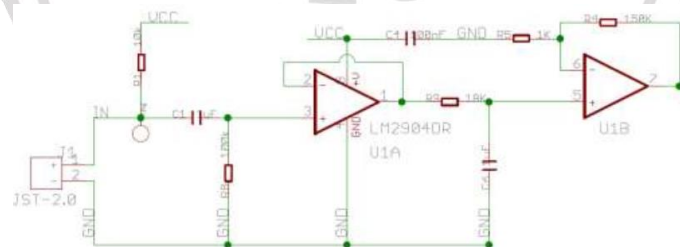
listrik hingga dikembangkan kepada penggunaan aplikatif seperti komparator dan osilator dengan distorsi rendah.

Penguat operasional dalam bentuk rangkaian terpadu memiliki karakteristik yang mendekati karakteristik penguat operasional ideal tanpa perlu memperhatikan apa yang terdapat di dalamnya. Karakteristik penguat operasional ideal adalah:

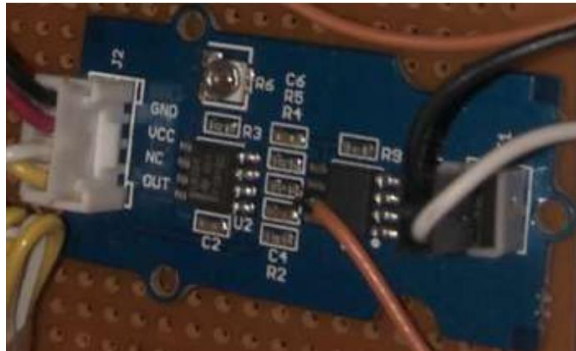
- Batas tegangan tidak terbatas.
- Impedansi masukan tidak terbatas.
- Impedansi keluaran nol.
- Lebar pita tidak terbatas.
- Tegangan offset nol (kondisi ketika masukan sebesar nol).

Gambar 2.2. di bawah ini merupakan skema penguat Inverting Amplifier sedangkan gambar 3.6. adalah rangkaian hardwarenya, menggunakan ic LM2904DR dan resistor $R_f = 1K$ Ohm, $R_i = 150K$ Ohm, maka di peroleh 150 kali penguatan.

$$A = -\frac{R_f}{R_i} = -\frac{1K}{150K} = -150 \text{ kali}$$



Gambar 2.2. Skema penguat Op – Amp

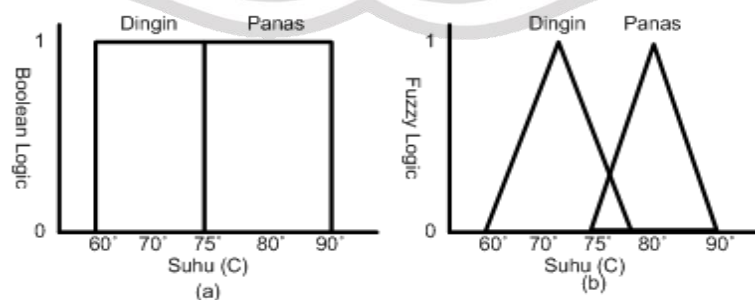


Gambar 2.3. Motherboard penguat Op – Amp

filter-filter HPF dan LPF dengan frekuensi cut off yang berbeda-beda digunakan untuk menganalisis sinyal pada skala yang berbeda. Sinyal dilewatkan pada sekumpulan HPF untuk menganalisis frekuensi tinggi dan dilewatkan pada LPF untuk menganalisis frekuensi rendah.

2.3. FUZZY LOGIC CONTROLLER

Sistem fuzzy ditemukan pertama kali oleh Prof. Lotfi Zadeh pada pertengahan tahun 1960 di Universitas California. Sistem fuzzy diciptakan karena boolean logic tidak mempunyai ketelitian yang tinggi, hanya mempunyai logika 0 dan 1 saja. Sehingga untuk membuat sistem yang mempunyai ketelitian yang tinggi tidak dapat menggunakan boolean logic. Perbedaan fuzzy logic dengan boolean logic terlihat pada gambar dibawah ini.



Gambar 2.4. Perbedaan *Boolean Logic* dan *Fuzzy Logic*

Pada gambar di atas, merupakan perbedaan antara logika boolean dengan logika fuzzy. Dimana logika fuzzy menggunakan parameter membership function sebagai indikator pendekatan pada kondisi nyata. Karena pada kenyataan di lapangan, parameter yang digunakan tidak selalu sama [20].

Kontroler logika fuzzy dikategorikan dalam kontrol cerdas (intelligent control). Unit logika fuzzy memiliki kemampuan menyelesaikan masalah perilaku sistem yang kompleks, yang tidak dimiliki oleh kontroler konvensional. Secara umum kontroler logika fuzzy memiliki kemampuan yaitu, Beroperasi tanpa campur tangan manusia secara langsung, tetapi memiliki efektifitas yang sama dengan kontroler manusia, Mampu menangani sistem-sistem yang kompleks, non-linier dan tidak stasioner, Memenuhi spesifikasi operasional dan kriteria kinerja dan Strukturnya sederhana, kokoh dan beroperasi real time.

Kebanyakan proses di dalam sistem pengendalian merupakan proses yang kompleks. Proses-proses tersebut pada kenyataannya bisa dikontrol secara manual dengan hasil yang cukup baik, oleh operator.

Operator tersebut adalah tenaga terampil yang mengandalkan pengalaman praktis, tanpa dilatarbelakangi teori-teori sistem pengendalian yang rumit. Pengalaman tersebut diperoleh dalam waktu yang lama, sehingga pengendalian yang dilakukan sepenuhnya bersifat intuitif. Untuk mendesain sistem pengendalian otomatis bagi proses – proses tersebut, diperlukan metode khusus, dalam hal ini digunakan teori logika fuzzy.

2.4. Transformasi Wavelet

Wavelet adalah fungsi matematika yang memotong-motong data menjadi kumpulan-kumpulan frekuensi yang berbeda, sehingga masing-masing komponen tersebut dapat dipelajari dengan menggunakan skala resolusi yang berbeda. Wavelet merupakan sebuah fungsi variabel real t , diberi notasi Ψt dalam ruang fungsi $L^2(\mathbb{R})$. Fungsi ini dihasilkan oleh parameter dilatasi dan translasi yang dinyatakan dalam Persamaan (2.1) [18-19]

$$\Psi_{a,b}(t) = a^{-1/2} \Psi\left(\frac{t-b}{a}\right); a > 0, b \in \mathbb{R} \quad (2.1)$$

Fungsi Wavelet berdasarkan Persamaan (2.1) terdiri dari a yang berarti parameter dilatasi dan b parameter translasi. Dengan mengambil $a = 2^{-j}$ dan $b = k2^{-j}$ [18-19], maka akan diperoleh sekumpulan fungsi basis yang ortogonal, artinya grafiknya tidak saling tumpang tindih.

$$\Psi_{j,k}(t) = 2^{j/2} \Psi(2^j t - k); j, k \in \mathbb{Z} \quad (2.2)$$

Sedangkan pada Persamaan Wavelet (2.2) [18-19], $2^{j/2}$ adalah parameter dilatasi dan k adalah parameter waktu atau lokasi ruang. Fungsi persamaan yang pertama dikenalkan pertama kali oleh Grossman dan Morlet, sedangkan persamaan yang kedua dikenalkan oleh Daubechies.

2.4.1. Discrete wavelet transform (DWT)

DWT berarti nilai yang digunakan dalam variabel-variabelnya adalah bilangan bulat positif. DWT menggunakan dua komponen penting dalam melakukan transformasi yakni fungsi skala (*scaling function*) dan fungsi wavelet (*wavelet function*). Fungsi skala disebut juga sebagai *Low Pass Filter* (LPF), sedangkan fungsi wavelet disebut juga sebagai

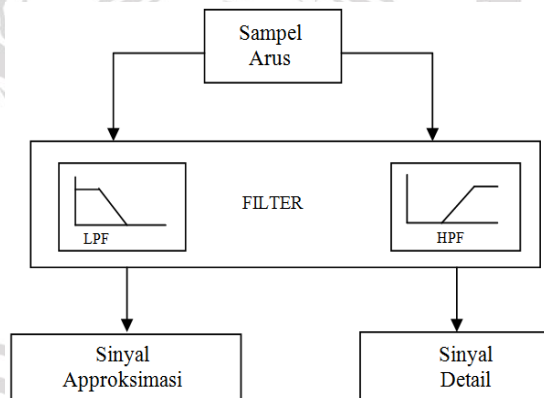
High Pass Filter (HPF). Kedua fungsi ini digunakan pada saat transformasi wavelet dan inverse transformasi wavelet.

Filter-filter HPF dan LPF dengan frekuensi cut off yang berbeda-beda digunakan untuk menganalisis sinyal pada skala yang berbeda. Sinyal dilewatkan pada sekumpulan HPF untuk menganalisis frekuensi tinggi dan dilewatkan pada LPF untuk menganalisis frekuensi rendah. Kedua komponen tersebut harus memenuhi kondisi seperti Persamaan (2.3).

$$\int_{-\infty}^{\infty} \Psi(x) dx = 0 \quad (2.3)$$

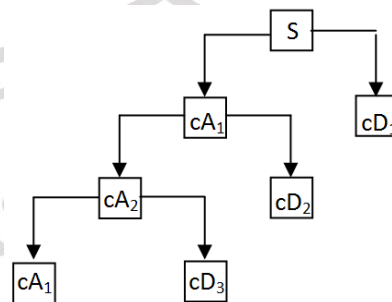
yang menjamin terjaminnya sifat ortogonalitas vektor.

Dalam analisis wavelet diskret dikenal istilah aproksimasi dan detail. Aproksimasi adalah komponen frekuensi rendah dan skala tinggi sinyal, sedangkan detail adalah komponen frekuensi tinggi dan skala rendah sinyal. Koefisien aproksimasi dan detail didapat melalui proses filter.



Gambar 2.5. Proses filter tingkat dasar

Proses filter pada tingkat dasar pada Gambar 2.5 dapat dijelaskan sebagai berikut. Sinyal pertama-tama dilewatkan pada rangkaian HPF dan LPF, kemudian setengah dari masing-masing keluaran diambil sebagai sampel melalui operasi pemfilteran. Proses ini disebut sebagai proses penguraian satu tingkat.



Gambar 2.6. Proses penguraian bertingkat tiga

Keluaran dari LPF digunakan sebagai masukan pada proses penguraian tingkat berikutnya seperti tampak pada Gambar 2.6. Proses ini diulang sampai tingkat proses penguraian yang diinginkan. Gabungan dari keluaran-keluaran HPF dan satu keluaran LPF yang terakhir, disebut sebagai koefisien wavelet, yang berisi informasi sinyal hasil transformasi yang telah terkompresi.

LPF dapat disebut juga dengan father wavelet dan HPF dapat disebut juga dengan mother wavelet. Hubungan antara mother wavelet dan father wavelet [18-19] adalah seperti pada Persamaan (2.4).

$$\Psi(t) = \Phi(2t) - \Phi(2t - 1) \quad (2.4)$$

Dengan $\Psi(t)$ adalah mother wavelet dan $\Phi(2t)$ adalah *father wavelet*. Fungsi persamaan dari *father wavelet* [13,14] adalah

$$\Phi_{j,k} = 2^{j/2} \Phi(2^j t - k), j, k \in \mathbb{Z} \quad (2.5)$$

Dari Persamaan (2.4) *father wavelet* dan Persamaan (2.5) *mother wavelet* tersebut akhirnya terbentuk persamaan transformasi wavelet secara lengkap. Persamaan transformasi wavelet tersebut biasanya ditulis dengan notasi sebagai berikut [13] :

$$f(t) = c_{0,0}\phi(t) + \sum_{j=0}^{M-1} \sum_{k=0}^{2^j-1} d_{j,k} \Psi_{j,k}(t) \quad (2.6)$$

Misalkan dalam sebuah peristiwa diambil 4 buah vektor pengamatan $x = (x_0, x_1, \dots, x_{p-1})^T$ dan dengan $p=2^M$ dan M adalah bilangan bulat positif, maka vektor pengamatan tersebut dinyatakan dalam fungsi tangga pada interval $[0,1]$. Hasilnya dapat dituliskan dalam bentuk matriks sebagai berikut [12-13]:

$$\begin{bmatrix} x_0 \\ x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} \phi\left(t \in \left[0, \frac{1}{4}\right)\right) & \Psi_{00}\left(t \in \left[0, \frac{1}{4}\right)\right) & \Psi_{10}\left(t \in \left[0, \frac{1}{4}\right)\right) & \Psi_{01}\left(t \in \left[0, \frac{1}{4}\right)\right) \\ \phi\left(t \in \left[\frac{1}{4}, \frac{2}{4}\right)\right) & \Psi_{00}\left(t \in \left[\frac{1}{4}, \frac{2}{4}\right)\right) & \Psi_{10}\left(t \in \left[\frac{1}{4}, \frac{2}{4}\right)\right) & \Psi_{01}\left(t \in \left[\frac{1}{4}, \frac{2}{4}\right)\right) \\ \phi\left(t \in \left[\frac{2}{4}, \frac{3}{4}\right)\right) & \Psi_{00}\left(t \in \left[\frac{2}{4}, \frac{3}{4}\right)\right) & \Psi_{10}\left(t \in \left[\frac{2}{4}, \frac{3}{4}\right)\right) & \Psi_{01}\left(t \in \left[\frac{2}{4}, \frac{3}{4}\right)\right) \\ \phi\left(t \in \left[\frac{3}{4}, 1\right)\right) & \Psi_{00}\left(t \in \left[\frac{3}{4}, 1\right)\right) & \Psi_{10}\left(t \in \left[\frac{3}{4}, 1\right)\right) & \Psi_{01}\left(t \in \left[\frac{3}{4}, 1\right)\right) \end{bmatrix} \begin{bmatrix} c_{00} \\ d_{00} \\ d_{10} \\ d_{11} \end{bmatrix}$$

2.5. Fuzzy Subspace Cluster

Fuzzy subspace cluster ialah pengelompokan berbasis jarak yang dimiliki kumpulan data, semakin jauh jarak antara titik data satu dengan titik data lainnya cenderung seragam. Ini tentu saja, menghambat efektivitas pengelompokan, karena pengelompokan berbasis jarak memanfaatkan bahwa jarak – jarak ini berbeda, dalam praktiknya seringkali hanya sebagian dari atribut yang tersedia

untuk membentuk ruang baru (*cluster*), meskipun ini mungkin tidak diketahui. Sebelumnya dalam kasus seperti ini, ingin memiliki algoritma pengelompokan yang secara otomatis menimbang data atau memilih bagian yang tepat.

Secara umum, ada tiga prinsip untuk melakukan seleksi fitur untuk membentuk ruang baru (*cluster*). Yang pertama adalah pendekatan filter pendekatan Iter (misalnya [8, 13]), yang mencoba menilai dan memilih fitur tanpa referensi eksplisit ke algoritma pengelompokan yang akan digunakan. Yang kedua adalah pendekatan pembungkus (misalnya [7, 9, 6]), yang menggunakan algoritma pengelompokan sebagai evaluator untuk subspacefitur yang dipilih dan dapat menggunakan strategi pencarian yang berbeda untuk memilih subspace yang akan dievaluasi. Pendekatan terakhir mencoba untuk menggabungkan clustering dan seleksi fitur dengan mendorong metode seleksi fitur ke dalam algoritmaclustering (misalnya [19, 17]). Perlu juga dicatat bahwa skema pembobotan fitur apa pun (yang mungkin, dengan sendirinya, menggunakan salah satu dari tiga prinsip ini) dapat diubah menjadi metode pemilihan fitur hanya dengan menerapkan ambang batas bobot ke bobot fitur yang dihitung.

Dalam makalah ini saya mempelajari pembobotan dan pemilihan fitur pada fuzzy clustering [1, 2, 12, 4]. Prinsip intinya adalah mentransfer ide alternatif ke fuzzy [15] ke atribut pembobotan fuzzyclustering [14]. Dengan memformulasi ulang Gustafson(Kesselfuzzyclustering [11], ini bahkan dapat diperluas ke skema untuk pembobotan dan pemilihan sumbu utama. Sementara dasar-dasar pendekatan ini telah diperkenalkan di [5] untuk pembobotan dan pemilihan atribut

global, makalah ini memperluas pendekatan ini untuk operasi khusus cluster. Dengan melakukan eksperimen pada data buatan maupun data dunia nyata, saya menunjukkan bahwa pendekatan ini bekerja dengan cukup baik dan mungkin sebenarnya sangat berguna dalam praktik, meskipun faktanya memerlukan fuzzy normal. pengelompokan yang dijalankan untuk inisialisasi (jika tidak, tidak cukup kuat) masih menyisakan ruang untuk perbaikan.

2.6. Pendahuluan dan notasi

Sepanjang makalah ini bahwa sebagai input set data m-dimensi \mathbf{X} yang terdiri dari n data $\mathbf{x}_j = (x_{j1}, \dots, x_{jm})$, $1 \leq j \leq n$. Kumpulan data ini juga dapat dilihat sebagai matriks data $\mathbf{X} = (x_{jk})_{1 \leq j \leq n, 1 \leq k \leq m}$, yang baris – barisnya merupakan titik data. Tujuannya adalah untuk mengelompokkan titik data ke dalam cluster, yang digambarkan oleh pusat cluster m-dimensi $\boldsymbol{\mu}_i = (\mu_{i1}, \dots, \mu_{im})$, $1 \leq i \leq c$. Pusat cluster ini serta bobot fitur yang akan diturunkan (karena dapat diartikan sebagai para meter bentuk dan ukuran cluster) secara bersama – sama dilambangkan dengan set parameter \mathbf{C} . Penetapan fuzzy dari titik data ke pusat cluster dijelaskan dengan matriks keanggotaan $\mathbf{U} = (u_{ij})_{1 \leq i \leq c, 1 \leq j \leq n}$.

2.7. Pembobotan data

Bagian ini mengulas dua metode dasar untuk menghitung bobot atribut dalam fuzzy clustering. Tujuannya utamanya adalah untuk membedakan metode – metode yang terkait erat ini untuk mengatur panggung untuk pendekatan pemilihan atribut yang dikembangkan dalam makalah ini.

2.7.1. Pengelompokan fuzzy axes-parallel gustafson-kessel

Cara yang sangat langsung untuk menentukan bobot atribut adalah dengan menerapkan fuzzy clustering axes-parallel Gustafson-Kessel[16]. Dalam hal ini kita harus meminimalkan fungsi tujuan.

$$J(\mathbf{X}, \mathbf{C}, \mathbf{U}) = \sum_{i=1}^c \sum_{j=1}^n h(u_{ij}) \sum_{k=1}^m \sigma_{i,k}^{-2} (x_{jk} - \mu_{ik})^2$$

Tunduk pada $\forall i, 1 \leq i \leq c : \prod_{k=1}^m \sigma_{i,k}^{-2} = 1$ dan kendala standar

$$\forall j, 1 < j \leq n : \sum_{i=1}^c u_{ij} = 1 \text{ dan } \forall i, 1 \leq i \leq c : \sum_{j=1}^n u_{ij} >$$

0. Varian $\sigma_{i,k}^{-2}$ yang diinginkan adalah spesifik cluster bobot atribut, yang harus ditemukan dengan mengoptimalkan fungsi transformasi keanggotaan h adalah fungsi cembung pada interval satuan.

2.7.2. Pengelompokan fuzzy bobot atribut

Sebuah alternatif, tetapi skema yang sama sederhananya untuk mendapatkan bobot atribut disarankan dalam [14]. Fungsi tujuan untuk meminimalkan adalah

$$J(\mathbf{X}, \mathbf{C}, \mathbf{U}) = \sum_{i=1}^c \sum_{j=1}^n h(u_{ij}) \sum_{k=1}^m w_{ik}^v (x_{jk} - \mu_{ik})^2.$$

Kendala pada derajat keanggotaan uji sama dengan fuzzy clustering Gustafson-Kessel, tetapi bobot atribut sekarang berbunyi

$$\forall i, 1 \leq i \leq c : \sum_{k=1}^m w_{ik} = 1. \text{ Parameter tambahan } v \text{ mengontrol}$$

pengaruh bobot atribut dengan cara yang sama seperti α (seperti dalam $h(u_{ij}) = u_{ij}^\alpha$) mengontrol pengaruh derajat keanggotaan. Aturan pembaharuan untuk derajat keanggotaan dan pusat cluster bertepatan dengan aturan fuzzy clustering Gustafson Kessel. Bobot diperbarui

sesuai dengan didefinisikan seperti pada bagian sebelumnya. Dengan menulis ulang aturan Gustafson-Kessel fuzzy clustering sebagai persamaan dan perbedaan menjadi sangat jelas : mereka terdiri dari normalisasi yang berbeda (jumlah bukan produk) dan parameter tambahan v .

2.8. Pemilihan atribut

Metode yang diluaskan diatas menghasilkan bobot atribut baik sebagian varian terbalik atau secara langsung sebagai bobot w_{ik} , $1 \leq i \leq c$, $1 \leq k \leq m$. Penting untuk dicatat bahwa dalam kedua kasus tidak mungkin bobot atribut apa pun hilang. Oleh karena itu modifikasi pendekatan diperlukan untuk memilih pada (yang dapat dicapai dengan memungkinkan bobot atribut menjadi 0).

Ide inti dari metode pemilihan atribut yang diusulkan adalah untuk menstransfer analisis efek α (seperti dalam $h(u_{ij}) = u_{ij}^\alpha$) dan alternatif yang mungkin, seperti yang dilakukan pada [15], ke bobot atribut. seperti yang ditunjukkan [15], perlu untuk menerapkan fungsi cembung $h(\cdot)$ ke derajat keanggotaan untuk mencapai penugasan fuzzy. Menaikkan derajat keanggotaan ke pangkat yang ditentukan pengguna, tentu saja seperti fungsi cembung, tetapi memiliki kelemahan bahwa memaksa semua penugasan menjadi kabur (yaitu, berbeda dari 0 dan 1). Alasannya adalah bahwa turunan dari fungsi ini hilang pada 0. Jika ingin mempertahankan kemungkinan penetapan tegas, kita lebih baik memilih fungsi h dengan $h'(0) > 0$.

2.9. Pembobotan sumbu utama

Masalah standar pendekatan pembobotan dan pemilihan atribut adalah bahwa atribut yang berkolerasi akan menerima bobot yang sangat mirip atau keduanya akan dipilih meskipun mereka jelas berlebihan. Untuk mengatasi masalah ini, pendekatan dalam semangat analisis komponen utama dapat digunakan. Alih – alih memberi bobot dan memilih atribut, dan dengan demikian (utama) sumbu kumpulan data. Bagian ini menunjukkan bagaimana metode bagian pembobotan data dapat diperluas ke pembobotan sumbu utama dengan memformulasi ulang pengelompokan fuzzy Gustafson-Kessel sehingga spesifikasi sumbu (utama) dan bobotnya dipisahkan.

2.9.1. Gustafson-Kessel fuzzy clustering

Gustafson-Kessel fuzzy clustering standar menggunakan jarak Mahalanobis (cluster-specific), yang didasarkan pada matriks kovarians spesifik cluster $\Sigma_i, i = 1, \dots, c$. Fungsi tujuannya adalah.

$$J(\mathbf{X}, \mathbf{C}, \mathbf{U}) = \sum_{i=1}^c \sum_{j=1}^n h(u_{ij})(\mathbf{x}_j - \boldsymbol{\mu}_i)^T \boldsymbol{\Sigma}_i^{-1} (\mathbf{x}_j - \boldsymbol{\mu}_i),$$

Yang akan diminimalkan dengan tunduk pada batasan.

$$\forall i, 1 \leq i \leq c: |\boldsymbol{\Sigma}_i^{-1}| = 1$$

(penafsiran intuitif volume cluster tetap) dan batasan standar

$$\forall j, 1 \leq j \leq n: \sum_{i=1}^c u_{ij} = 1 \text{ dan } \forall i, 1 \leq i \leq c: \sum_{j=1}^n u_{ij} > 0.$$

aturan pembaruan yang dihasilkan untuk matriks kovarians adalah Σ_i .

Untuk mendapatkan bobot eksplisit dari sumbu (utama), kita amati bahwa Σ_i karena merupakan matriks definit simetris dan positif, maka matriks tersebut memiliki dekomposisi nilai eigen $\Sigma_i = \sigma_{i,1}, \dots, \sigma_{i,m}$

dengan $\mathbf{D}_i = \text{diag}(\mathbf{R}_i \mathbf{D}_i^2 \mathbf{R}_i^\top)$ (i.e., nilai eigen $\sigma_{i,1}^2$ ke $\sigma_{i,m}^2$) dan matriks ortogonal \mathbf{R}_i , yang kolom-kolomnya merupakan vektor eigen yang bersesuaian. ini memungkinkan kita untuk menulis invers matriks kovarians Σ_i sebagai dengan. sebagai konsekuensinya, kita $\Sigma_i^{-1} = \mathbf{T}_i \mathbf{T}_i^\top$ ulang fungsi tujuan sebagai $\mathbf{T}_i = \mathbf{R}_i \mathbf{D}_i^{-1}$

$$\begin{aligned} J(\mathbf{X}, \mathbf{C}, \mathbf{U}) &= \sum_{i=1}^c \sum_{j=1}^n h(u_{ij}) (\mathbf{x}_j - \boldsymbol{\mu}_i)^\top \mathbf{T}_i \mathbf{T}_i^\top (\mathbf{x}_j - \boldsymbol{\mu}_i) \\ &= \sum_{i=1}^c \sum_{j=1}^n h(u_{ij}) ((\mathbf{x}_j - \boldsymbol{\mu}_i)^\top \mathbf{R}_i \mathbf{D}_i^{-1}) ((\mathbf{x}_j - \boldsymbol{\mu}_i)^\top \mathbf{R}_i \mathbf{D}_i^{-1})^\top \\ &\quad - \sum_{i=1}^c \sum_{j=1}^n h(u_{ij}) \sum_{k=1}^m \sigma_{i,k}^{-2} \left(\sum_{l=1}^m (x_{jl} - \mu_{il}) r_{i,lk} \right)^2. \end{aligned}$$

dalam bentuk ini penskalaan dan rotasi ruang data yang dicakup dalam matriks kovarians dipisahkan dengan baik, yang pertama diwakili oleh varian $\sigma_{i,k}^2$, $k = 1, \dots, m$ (atau kebalikannya), yang terakhir oleh matriks ortogonal \mathbf{R}_i . memberikan bobot sumbu yang diinginkan, sedangkan vektor eigen yang sesuai (kolom atau \mathbf{R}_i) menunjukkan sumbu (utama).

2.9.2. Reformulasi dari gustafson-kessel fuzzy clustering

Untuk mentransfer pendekatan [14] dan yang dikembangkan di Bagian 4, kita mulai dari fungsi tujuan yang ditulis ulang, di mana penskalaan dan rotasi ruang data dipisahkan dan dengan demikian dapat diperlakukan secara independen. menurunkan aturan pembaruan untuk faktor penskalaan $\sigma_{i,k}^{-2}$ adalah trival, karena pada dasarnya hasil yang sama diperoleh untuk pengelompokan fuzzy gustafson kessel sumbu paralel.

$$\sigma_{i,k}^2 = s_{i,k}^2 \left(\prod_{r=1}^m s_{i,r}^2 \right)^{-\frac{1}{m}}$$

2.10. Pemilihan Sumbu Utama

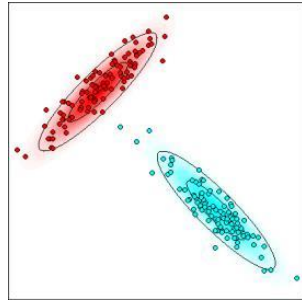
Dalam analogi transisi dari pembobotan atribut ke pemilihan atribut, transisi dari pembobotan sumbu utama ke pemilihan sumbu utama : kita cukup mengganti aturan pembaruan untuk bobot (yang sekarang terpisah dari sumbu) dengan yang diperoleh di bagian pemilihan atribut. Ini mengarah ke aturan pembaruan.

$$\forall i; 1 \leq i \leq c : \forall k; 1 \leq k \leq m : w_{ik} = \frac{1}{1-\beta} \left(\frac{1 + \beta(m_{i\oplus} - 1)}{\sum_{r=1; w_{ir} > 0}^m s_{i,r}^{-2}} s_{i,k}^{-2} - \beta \right)$$

Dengan, didefinisikan seperti dalam bagian reformulasi dari gustafson-kessel fuzzy clustering dan sebagaimana didefinisikan dalam bagian pemilihan atribut.

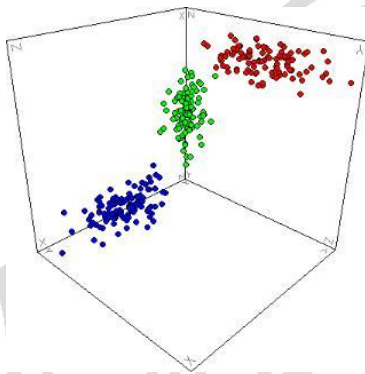
2.11. Eksperimen

Dari semua eksperimen yang saya lakukan dengan metode yang dijelaskan pada berbagai kumpulan data, saya melaporkan hanya beberapa di sini, karena keterbatasan ruang. Karena hasil eksperimen untuk pembobotan global dan pemilihan atribut dapat ditemukan di [5], saya membatasi diri di sini pada pembobotan dan pemilihan atribut khusus cluster.



$\epsilon = 0:5$	X	Y	U_1	U_2	U_3
r_1	0:66	0:76	0:02	0:02	0:00
r_2	0:67	0:74	0:00	0:02	0:01

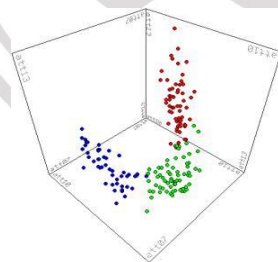
Gambar 2.7. Menunjukkan dua kumpulan data



$\epsilon = 0:3$	X	Y	Z	U_1	U_2	U_3
$w_{1;k}$	0.63	0.37				
$w_{2;k}$	0.13		0.87			
$w_{3;k}$		0.29	0.71			

$\epsilon = 0:4$	X	Y	Z	U_1	U_2	U_3
$w_{1;k}$	1.00					
$w_{2;k}$			1.00			
$w_{3;k}$		0.28	0.72			

Gambar 2.8. Kumpulan data artifisial dengan tiga ruang gaussian dan 300 titik data



$\epsilon = 0:3$	att7	att10	att13
$w_{1;k}$			1.00
$w_{2;k}$		1.00	
$w_{3;k}$	0.77	0.23	

Gambar 2.9. Kumpulan data dunia nyata dengan tiga kelas dan 178 titik data diagram menunjukkan bagian banyaknya data

Gambar 2.7, 2.8 dan 2.9 menunjukkan dua kumpulan data buatan dan satu dunia nyata dan hasil pengelompokan yang diperoleh dari mereka. Dalam ketiga kasus, algoritma diinisialisasi dengan pengelompokan fuzzy Gustafson-Kessel sumbu-paralel (lihat Bagian gambar 2.9), yang dijalankan hingga konvergensi. (Tanpa inisialisasi seperti itu, hasilnya tidak cukup stabil.) Seperti yang dapat dilihat dari hasil ini, metode ini menjanjikan dan mungkin sebenarnya sangat berguna dalam praktik. Dalam ketiga kasus, atribut yang tidak informatif dihilangkan dengan baik (menerima bobot nol atau koefisien mendekati nol). Sedangkan atribut informatif menerima bobot tinggi, yang dengan baik mencerminkan struktur kumpulan data.

2.12. Transformasi Wavelet Haar

Wavelet haar merupakan salah satu tipe wavelet yang paling sederhana yang dapat diterapkan pada transformasi signal dan transformasi pada citra.

2.13. Transformasi signal

Fungsi basis ruang V_j disebut dengan fungsi skala dan disimbolkan sebagai ϕ . Salah satu fungsi basis sederhana dinyatakan sebagai berikut:

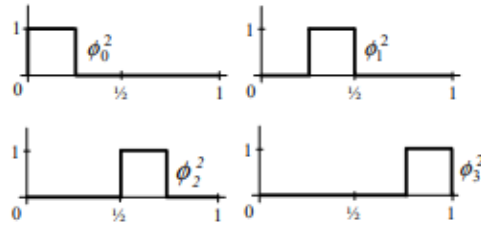
$$\phi_i^j(x) := \phi(2^j x - i), \quad i = 0, \dots, 2^j - 1 \quad (1)$$

Dengan

$$\Phi(x) := \begin{cases} 1, & \text{untuk } 0 \leq x \leq 1 \\ 0, & \text{untuk } x \text{ yang lainnya} \end{cases}$$

Fungsi diatas disebut juga sebagai fungsi kotak (box function). Sebagai contoh fungsi kotak pembentuk basis dalam ruang V^2 akan terdapat 2^j atau $2^2 = 4$

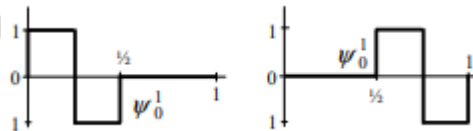
potongan fungsi konstan pada jangkauan [0,1] seperti terlihat pada Gambar 2.7 berikut ini.



Gambar 2.10. Fungsi wavelet haar pada dalam kotak

Fungsi wavelet yang sesuai dengan fungsi penyekalaan diatas disebut dengan wavelet haar, yang diberikan dengan persamaan.

$$\psi_i^j(x) = \psi(2^j x - i), i=0, 1, \dots, 2^j-1 \quad (2)$$



Gambar 2.11. Fungsi wavelet haar pada luar kotak