

PEMBANGKITAN ANIMASI WAJAH KARAKTER *LIFE-LIKE* UNTUK EKSPRESI EMOSI

Surya Sumpeno

Jurusan Teknik Elektro, Fakultas Teknologi Industri
Institut Teknologi Sepuluh Nopember (ITS) Surabaya
E-mail: surya@ee.its.ac.id

Abstract

A model consists of text classifier and Mamdani Fuzzy Inference System to build a life-like virtual character capable of expressing emotion from a text input has been proposed. In real world application such as animation movie, the life-like virtual character of proposed model needs to be animated.

This paper describes animation characteristics of the model using neutral expression as centre of facial expression transition from one emotion to another. Emotion transition can be viewed as gradual decrease or increase of emotion intensity from one emotion toward other emotion. Experimental results show that animations of life-like character can be generated automatically using proposed model. However, future work is needed in order to maintain animation coherence as workable and possible solutions for drastic change of value in controller signals.

Keywords: Animation of Life-like Character, Facial Expression of Emotion.

1. PENDAHULUAN

Teks tidak hanya memuat informasi, tetapi mampu memicu tanggapan emosional dari pembaca (pendengar) atau penulis (pembicara). Contoh, jika seseorang membaca judul artikel surat kabar “Pesawat Terbang Berpenumpang 51 Orang, Jatuh 36 Selamat”¹, maka boleh jadi ia merasa sedih dan barangkali pula takut.

Respon yang muncul dari teks contoh ini dapat berupa ekspresi emosi yang teramati dalam wujud verbal dan non-verbal. Tujuan makalah ini adalah membangun karakter yang hidup (*life-like*) yang mampu merespon emosi dari teks masukan, secara non verbal, karena ekspresi emosi yang non verbal sangat vital

dalam pembentukan karakter yang *life-like*. Menurut [1] respon emosional merupakan salah satu komponen pokok dalam karakter *life-like*, yaitu karakter yang menyajikan ilusi kepada pemirsa bahwa ia adalah karakter yang hidup. Fokus dalam makalah ini adalah perilaku non verbal karakter dalam bentuk ekspresi wajah.

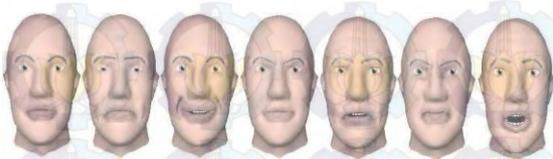
Untuk membuat karakter yang dapat dipercaya sebagai karakter yang hidup, para animator --khususnya lagi dari studio Disney-- telah memasukkan unsur emosi ke dalam animasi karakter untuk membuat pemirsa percaya akan hidupnya karakter tersebut, dalam petualangan dan ketidak-beruntungan yang membuat pemirsa bisa tertawa dan menangis [2]. Para animator yakin bahwa emosi bila diekspresikan pada saat yang tepat adalah kunci dalam membuat animasi karakter yang berkualitas dan hidup (*life-like*).

Sebuah model yang mencoba mengimplementasikan karakter virtual yang mampu mengekspresikan emosi dalam bentuk tampilan wajah sesaat (*snapshot*), setelah menerima masukan teks, telah diajukan oleh [3]. Model ini mendasarkan pada asumsi-asumsi bahwa emosi campuran tersusun atas emosi-emosi dasar [4] dan setiap emosi dasar dapat memiliki ekspresi wajah yang berbeda [5] dari ekspresi emosi dasar lainnya. Dengan demikian, ekspresi wajah yang tercampur merupakan hasil kombinasi dari ekspresi-ekspresi wajah untuk emosi-emosi dasar.

Pada penelitian yang lain [6] telah merancang sebuah model karakter dengan enam (6) ekspresi wajah untuk emosi dasar seperti ditunjukkan pada Gambar 1. Sementara paparan tekstual dari karakteristik ekspresi-ekspresi wajah sebagai representasi dari emosi-emosi dasar, dapat mengacu pada [7].

Sistematika penulisan adalah sbb: di bagian metode akan dijelaskan lebih lanjut tentang model yang mampu membangkitkan ekspresi wajah sesaat, yang terkait dengan nilai emosi yang mungkin terkandung dalam sebuah teks dan bagaimana membangkitkan animasinya.

¹ <http://technews.tmcnet.com/topics/associated-press/articles/102174-plane-carrying-51-crashes-venezuela-36-survive.htm> published at: September 16, 2010



Gambar 1. Salah satu contoh karakter dengan ekspresi emosi pada wajah [6].

Sedangkan di bagian diskusi akan ditampilkan hasil-hasil eksperimen dan penjelasannya. Pada bagian hasil, diberikan simpulan dari makalah ini.

2. METODE

Model terdiri dari dua modul yaitu (i) pengklasifikasi teks berbasis probabilitas Bayesian dan (ii) sistem inferensi berbasis *fuzzy* untuk pengendali ekspresi wajah karakter virtual, sebagaimana dapat dilihat pada Gambar 2.

Nilai-nilai probabilitas dari kelas-kelas emosi dasar yang merupakan keluaran dari modul pengklasifikasi teks Bayesian diumpangkan ke sistem inferensi *fuzzy* yang mengendalikan “otot-otot” penggerak wajah karakter virtual. Diasumsikan bahwa nilai-nilai keluaran tersebut merupakan tingkat kekuatan suatu emosi dasar dirasakan.

Misalnya, bila probabilitas nilai “senang” (P_{joy}) sama dengan 1.0 (yang merupakan nilai maksimum sebuah probabilitas), artinya emosi

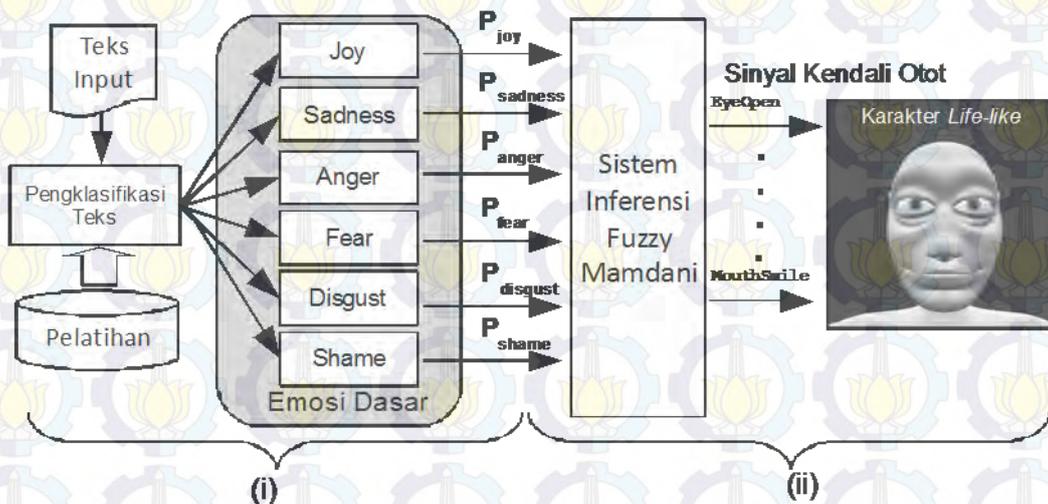
yang dirasakan oleh karakter adalah senang maksimum, sedangkan nilai $P_{joy}=0.0$ berarti emosi dasar senang tidak berdampak pada ekspresi wajah.

Pengklasifikasi teks Naive Bayes diimplementasikan memanfaatkan Weka [8], sedangkan karakter virtual *life-like* menggunakan karakter Ludwig [a] berlisensi *common creative* yang berbasis Blender [b] sebuah aplikasi grafis 3D sumber terbuka dan mendayagunakan bahasa pemrograman *script* Python [c].

Pengendali *fuzzy* diimplementasikan menggunakan perangkat lunak sumber terbuka *fuzzyLogic* [d] yang berbasis FCL (Fuzzy Control Language) [9]. Dengan mengadopsi format standar FCL, bahasa pemrograman tidak lagi menjadi isu. Bahasa python pun dapat dipakai memanfaatkan *pyfuzzy* [e].

Sistem inferensi *fuzzy* yang dipakai adalah inferensi *fuzzy* Mamdani. Logika *fuzzy* dimanfaatkan karena dipandang cocok bagi model yang perlu melakukan kalkulasi terhadap keadaan “ragu-ragu” seperti “sedikit”, “cukup” dan “agak”.

Untuk dapat menganimasikan perubahan ekspresi dari satu emosi ke emosi yang lain, model mesti dilengkapi dengan penambahan satu titik acuan yang menggambarkan ekspresi wajah netral, yaitu ekspresi wajah tanpa emosi. Baik [3] maupun [6] sama-sama menggunakan karakter dengan senyuman sangat tipis sebagai ekspresi wajah netral.

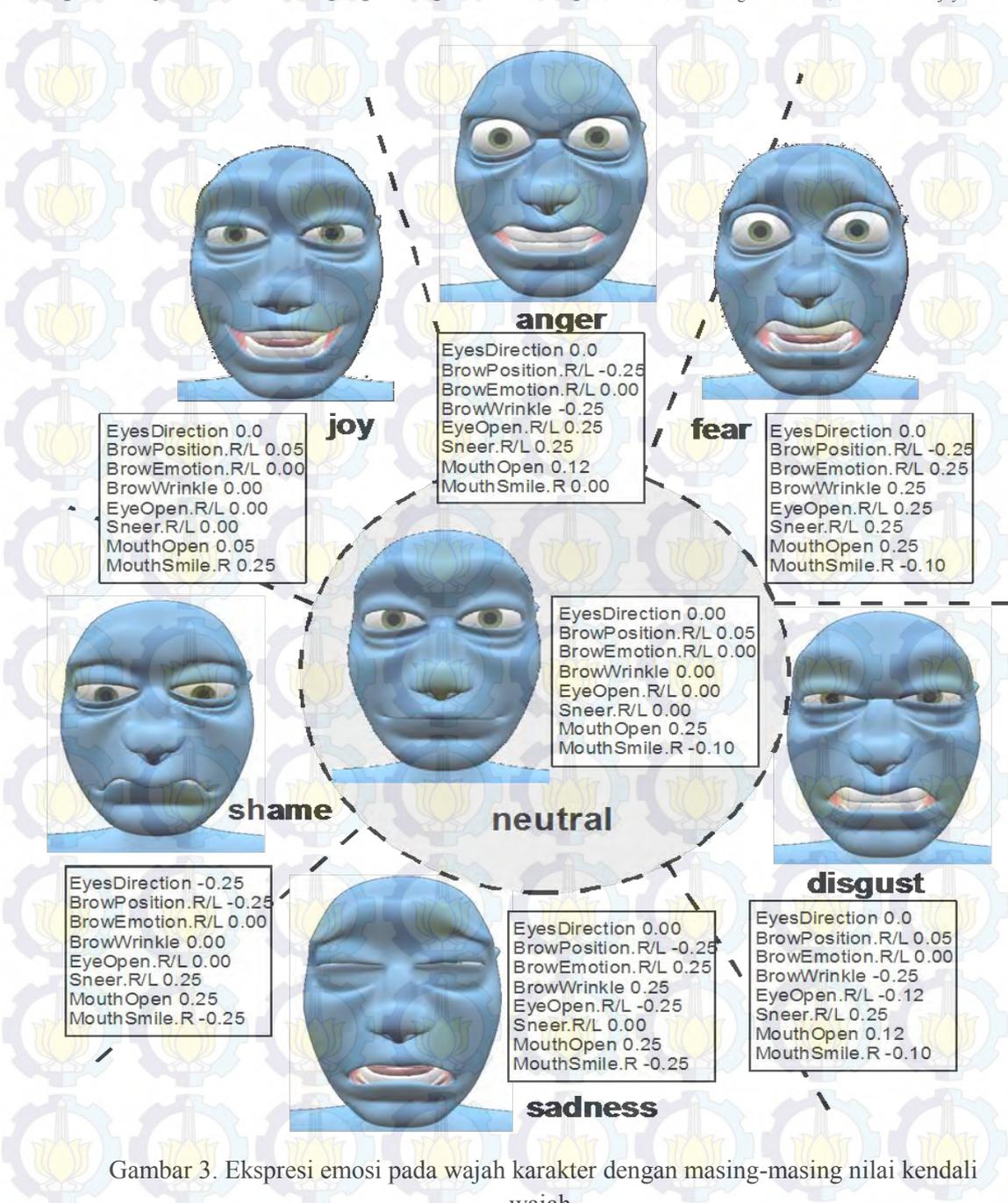


Gambar 2. Sistem pembangkit ekspresi wajah karakter life-like terkait emosi dalam teks.

Perubahan dari satu jenis emosi ke jenis emosi yang lain digambarkan sebagai transisi satu jenis emosi, ke arah tanpa emosi (netral) kemudian berpindah menuju ke emosi yang lain tersebut. Contoh, perubahan emosi dari “agak senang” ke “sedikit marah” diwujudkan dalam transisi dari “agak senang” ke “netral” baru dilanjutkan ke “sedikit marah”. Mekanisme dengan netral sebagai pusat ini diilustrasikan pada Gambar 3. Tampak netral di tengahnya dikelilingi oleh emosi-emosi dasar yang berekspresi wajah terkait dilengkapi dengan ni-

lai-nilai variabel-variabel sinyal kendali wajah karakter Ludwig.

Perubahan ekspresi wajah karakter dilakukan dengan melakukan perubahan pada nilai-nilai masukan sistem inferensi *fuzzy* di modul (ii), yang berupa nilai-nilai tingkat kekuatan dari setiap emosi dasar secara berangsur-angsur. Jika diambil contoh transisi emosi “agak senang” (yang misalnya diwakili oleh nilai probabilitas $P_{joy} = 0.7$) ke “sedikit marah” (yang misalnya diwakili oleh nilai probabilitas $P_{anger} = 0.2$), maka P_{joy} berangsur



Gambar 3. Ekspresi emosi pada wajah karakter dengan masing-masing nilai kendali wajah.

turun ke netral ($P_{joy} = 0.0$), kemudian diikuti dengan P_{anger} yang bergerak mulai dari 0.0 berangsur naik ke angka 0.2.

Nilai penambahan/pengurangan dari transisi emosi ini tergantung pada berapa banyak gambar (*frame*) yang akan dibangkitkan dalam satu detik (*frames-per-second/fps*). Saat ini laju *frame* untuk tayangan video bervariasi mengikuti standar *broadcast* televisi, yaitu 24 fps, 25 fps, dan 30 fps masing-masing untuk PAL/SECAM dan NTSC. Standar yang lebih modern seperti HDTV ATCS (Advanced Television Systems Committee) [f] membutuhkan laju frame yang lebih tinggi sebesar 60 fps, bahkan untuk tayangan film 3D bisa mencapai 120 fps.

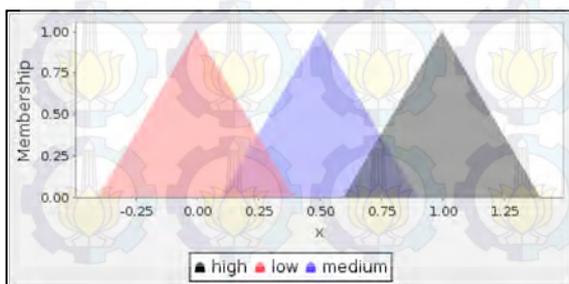
Dalam makalah ini, yang dibahas cukup detail adalah grafik fungsi masukan dan keluaran dari modul pengendali *fuzzy*, karena terkait langsung dengan pembangkitan animasi dari karakter. Untuk pembahasan lebih rinci tentang model selengkapnya, yang meliputi pula kinerja pengklasifikasi teks, dapat mengacu pada [3].

Sistem inferensi *fuzzy* Mamdani tersusun atas 3 tahapan, yang mencakup fungsi keanggotaan *fuzzy* masukan, pemetaan berbasis aturan dan defuzifikasi keluaran.

Fungsi keanggotaan *fuzzy* masukan terdiri dari tiga variabel bahasa, yaitu “rendah”, “sedang” dan “tinggi” untuk masing-masing nilai emosi dasar (*joy*/senang, *fear*/takut, *anger*/marah, *sadness*/sedih, *disgust*/jijik dan *shame*/malu) diimplementasikan dengan fungsi segitiga sebagaimana dapat dilihat pada Gambar 4.

Prinsip desain defuzifikasi pada model ini adalah memasang logika “1” (nilai maksimum keluaran sinyal kendali) ke semua variabel yang termuat dalam ekspresi dasar (pada Gambar 3) ke nilai masing-masing, sedangkan pergeseran ke arah-arah yang lain, menuju ke posisi ekspresi wajah netral.

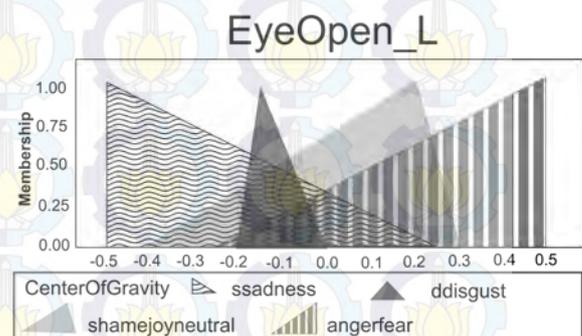
Aturan pemetaan disusun berdasar atas beberapa karakteristik antara lain:



Gambar 4. Fungsi keanggotaan masukan.

- Nilai total probabilitas dari masukan tidak pernah melebihi angka 1.0. Dengan demikian, tidak mungkin semua masukan bernilai “tinggi”. Maksimum hanya satu variabel masukan yang dapat bernilai “tinggi”.
- Pemetaan dilakukan pada sinyal-sinyal kendali yang mengalami perubahan yang dikaitkan dengan emosi tertentu. Contoh, variabel sinyal kendali *EyesDirection* hanya dipengaruhi oleh nilai emosi “malu”.
- Bila semua nilai masukan bernilai rendah, maka defuzifikasi diarahkan ke ekspresi “netral”.
- Memanfaatkan pembobotan (*WITH nilai_bobot*), aturan untuk nilai masukan “tinggi” dapat diterapkan pada nilai masukan lebih rendah, dengan menetapkan *nilai_bobot* yang lebih rendah pula.
- Defuzifikasi menggunakan perhitungan *Center of Gravity* (COG) dan fungsi segitiga.

Hasilnya, jumlah total aturan ada 49, yang meliputi kendali *EyesDirection* 2 aturan, *BrowPosition* 5 aturan, *BrowEmotion* 5 aturan, *BrowWrinkle* 7 aturan, *EyeOpen* 9 aturan, *Sneer* 5 aturan, *MouthOpen* 7 aturan dan *MouthSmile* 9 aturan. Sebagai contoh implementasi, Tabel 1 dan Gambar 5 memperlihatkan aturan dan defuzifikasi untuk sinyal kendali *EyeOpen*.



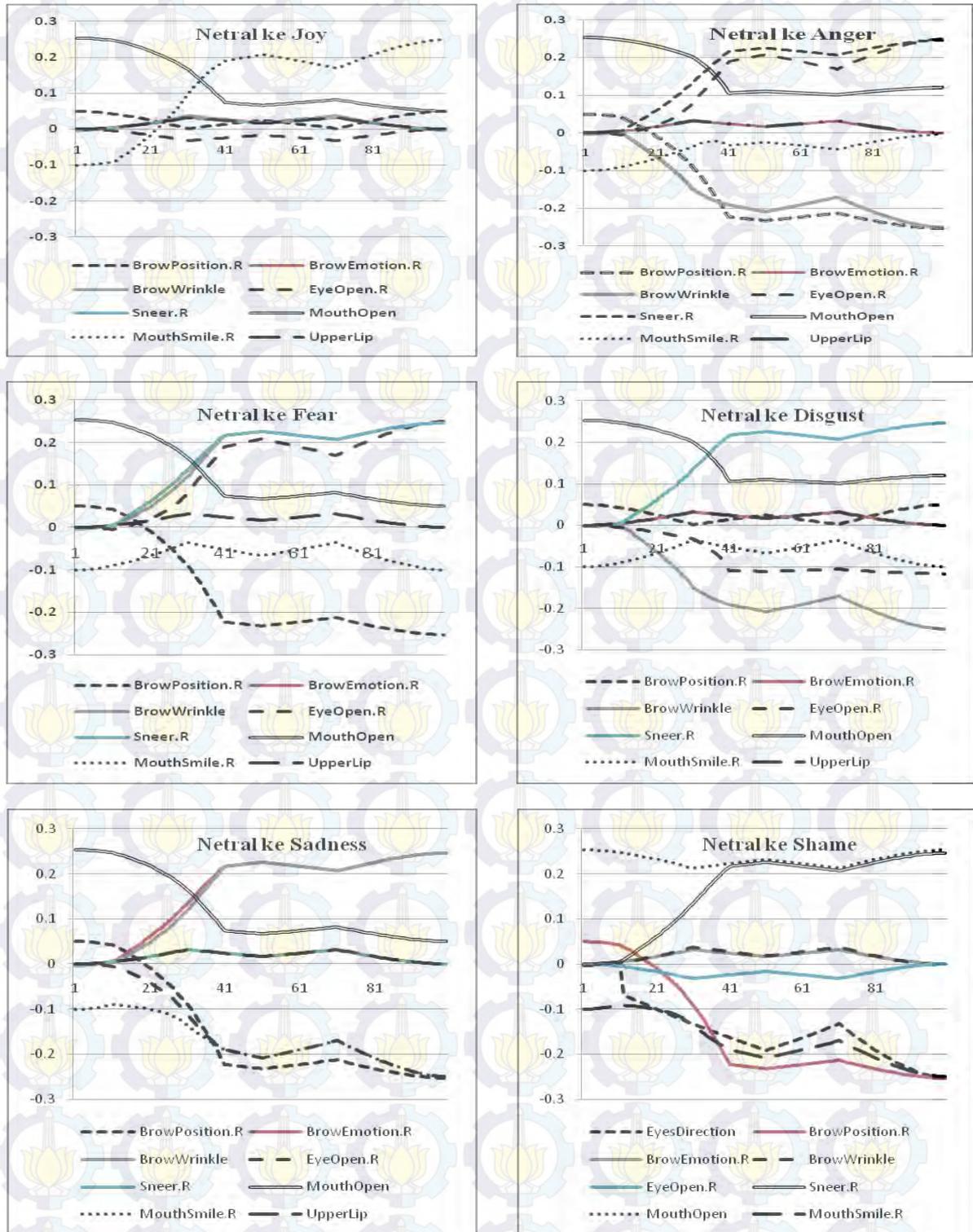
Gambar 5. Defuzifikasi untuk sinyal kendali *EyeOpen*

3. DISKUSI

Untuk melihat pergerakan nilai-nilai sinyal kendali ekspresi wajah secara lengkap, nilai-nilai variabel masukan ditetapkan bergerak turun dari posisi puncak (misalnya $P_{joy} = 1.0$) berangsur menuju netral. Gambar 6 memperlihatkan perubahan nilai-nilai kendali eks-

presi wajah tersebut dari posisi netral ke masing-masing posisi puncak emosi dasar. Grafik pergerakan dari posisi puncak masing-

masing emosi dasar ke netral tidak diperlihatkan, karena hanya fungsi kebalikannya. Terlihat beberapa patahan di beberapa titik,



Gambar 6. Perubahan nilai-nilai kendali ekspresi wajah karakter, ketika mengalami perpindahan emosi dari netral.

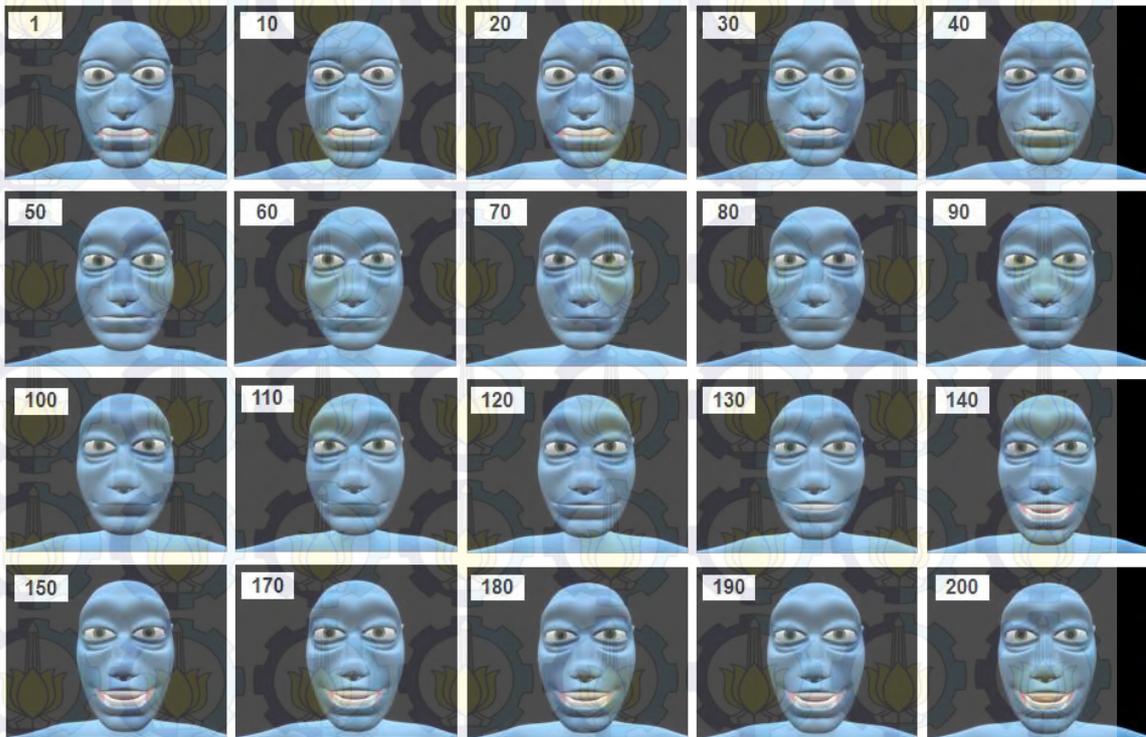
tetapi asalkan patahannya tidak sangat tajam, hasil animasi ekspresi wajah dalam pengamatan visual eksperimen, masih dapat dilihat terkesan mulus (*smooth*), tidak terpatah. Namun memang di patahan ekstrim seperti misalnya pada sinyal kendali *EyesDirection* untuk perubahan emosi dari netral ke *shame*, animasi pergerakan mata menjadi tidak mulus karena adanya penurunan drastis yang sangat curam di sekitar posisi frame (lihat di segmen 1~21). Ini memerlukan perbaikan berupa *adjustment* dan penghalusan (*refinement*) di kerja lanjutan.

Gambar 7 memperlihatkan cuplikan anima-

si untuk perpindahan ekspresi emosi karakter virtual dari mendapat masukan teks berupa kalimat pertama “*When friends try to put me down or hurt me*” (yang memiliki nilai keluaran dari pengklasifikasi teks $P_{\text{anger}}=0.431$, $P_{\text{sadness}}=0.265$, $P_{\text{shame}}=0.232$, $P_{\text{disgust}}=0.042$, $P_{\text{fear}}=0.028$ dan $P_{\text{joy}}=0.001$), menuju ke ekspresi emosi dari masukan kalimat kedua “*When I met friends I had not seen for the last 2-4 years*” yang memiliki nilai keluaran pengklasifikasi teks $P_{\text{joy}}=0.915$, $P_{\text{sadness}}=0.048$, $P_{\text{shame}}=0.022$, $P_{\text{fear}}=0.009$ dan $P_{\text{anger}}=P_{\text{disgust}}=0.003$). Aplikasi dari contoh-contoh kalimat ini adalah untuk

Tabel 1. Aturan untuk kendali *EyeOpen*

RULE 1: IF (disgust IS low) AND (fear IS low) AND (joy IS low) AND (sadness IS low) AND (shame IS low) AND (anger IS low) THEN EyeOpen L IS shamejoyneutral;
RULE 2: IF (fear IS high) OR (anger IS high) THEN EyeOpen L IS angerfear;
RULE 3: IF (joy IS high) OR (shame IS high) THEN EyeOpen L IS shamejoyneutral;
RULE 4: IF (sadness IS high) THEN EyeOpen L IS ssadness;
RULE 5: IF (disgust IS high) THEN EyeOpen L IS ddisgust;
RULE 6: IF (sadness IS medium) THEN EyeOpen L IS ssadness WITH 0.5;
RULE 7: IF (joy IS medium) OR (shame IS medium) THEN EyeOpen L IS shamejoyneutral WITH 0.5;
RULE 8: IF (anger IS medium) OR (fear IS medium) THEN EyeOpen L IS angerfear WITH 0.5;
RULE 9: IF (disgust IS medium) THEN EyeOpen L IS ddisgust with 0.5;



Gambar 7. Cuplikan animasi ekspresi emosi karakter virtual dari mendapat masukan teks “*When friends try to put me down or hurt me*” menuju ke ekspresi “*When I met friends I had not seen for the last 2-4 years.*”

karakter *life-like* pembaca cerita (*virtual story teller*) atau teman virtual yang mendengarkan celotehan (atau membaca teks) tentang kesan-kesan suka-duka pertemanan.

4. HASIL DAN KERJA LANJUTAN

Model yang mampu membangkitkan animasi secara otomatis untuk ekspresi wajah beremosi dari masukan teks, demi membangun sebuah karakter yang hidup (*life-like*) telah ditunjukkan, dengan hasil yang cukup memuaskan, namun di beberapa aspek masih memerlukan penyempurnaan yang akan dilakukan di kerja lanjutan mendatang.

5. DAFTAR PUSTAKA

- [1] Joseph Bates. The role of emotion in believable agents. *Communications of the ACM*, 37:122-125, 1994.
- [2] Thomas, F. and Johnston, O. *The Illusion of Life*. Abbeville Press, New York, 1981.
- [3] Surya Sumpeno dkk. Text Mining for Fuzzy-based Emotion Expression, *The Journal for Technology and Science (IPTEK) vol. 21 No. 1*, ISSN 0853-4098, February 2010.
- [4] Robert Plutchik. *The Emotions*. University Press of America, Inc., July 25, 1991
- [5] Paul Ekman. Facial expression and emotion. *American Psychologist*, 48:384-0392, 1993.
- [6] The Duy Bui. Creating emotions and facial expressions for embodied agents. PhD thesis, Enschede, Netherlands, July 2004. URL <http://doc.utwente.nl/41553/>.
- [7] A. Murat Tekalp and Jrn Ostermann. Face and 2-d mesh animation in mpeg-4. *Signal Processing: Image Communication*, 15(4-5):387-421, 2000. ISSN 0923-5965
- [8] M. Hall, E. Frank, G. Holmes, B. Pfahringer, P. Reutemann and I. H. Witten. The weka data mining software: an update. *SIGKDD Explor. Newsletter*, vol. 11, no. 1, pp. 10–18, 2009
- [9] Technical Committee No.65, *IEC 1131-Programmable Controllers*, International Electrotechnical Commission (IEC), 1997.

Website

- [a] <http://jasonpierce.animadillo.com/resource/ludwig/ludwig.html>
- [b] <http://www.blender.org/>
- [c] <http://www.python.org/>
- [d] <http://jfuzzylogic.sourceforge.net/>
- [e] <http://pyfuzzy.sourceforge.net/>
- [f] <http://en.wikipedia.org/wiki/ATSC>