

Automatinis veido išraiškos atpažinimas

D. Dervinis

Šiaulių universitetas,

Vilniaus g. 141, LT – 76353 Šiauliai, Lietuva, el. p. ddonatas@cr.su.lt

Įvadas

Kuriant žmogaus ir mašinos sistemas, vienas iš sąveikos būdų yra galvos judesiai. Valdant galva, svarbu įvertinti pasirinktų galvos ir veido taškų pokyčius. Registruodami galvos judesius dėl veido išraiškos kitimo galime gauti klaidingus rezultatus – registruojamo taško koordinatės kinta ne tik sukant ar lenkiant galvą, bet ir keičiantis veido išraiškai.

Sukaupiti taškų pokyčių statistiniai duomenys, kai keičiasi emocijos, leidžia nustatyti galvos posūkio kampa panaikinant įvairių išraiškų padarytas paklaidas. Taip pat kiekybiškai nustačius ir įvertinus emocijas, galima kurti pažangias kalbos atpažinimo ar sintetinio sistemas. Pagal operatoriaus emocijas reikėtų vertinti ir pritaikyti darbo vietą. Automatizuotas veido išraiškų nustatymas gali tapti patogus psichologų ar terapeutų įrankis, kuris padėtų vienareikšmiškai ir kiekybiškai įvertinti žmogaus emocijų būklę.

Daugelis tyrėjų veido išraiškas nustato kadrų sekoje naudodami taškų poslinkius [1] arba veido asimetriškumo savybes [2]. Šie būdai netinka, jei turime tik statinį vaizdą su viena veido išraiška. Ankstesni tyrimai rodo, kad būdingi taškai, naudojami galvos posūkiui nustatyti statiniame vaizde, išsidėstę apie akis, antakius ir lūpas [3]. Tai taškai kurie kinta, veikiami išraiškų.

Darbe analizuojami taškų koordinatinių poslinkiai bei taškų tarpusavio pokyčiai, kurie leis pagal pirminį (atraminį arba neutralios išraiškos) ir galutinį kadrus nustatyti veido išraišką. Pasiūlytu būdu identifikuojama viena iš 5 pagrindinių veido išraiškų, taip pat kiekybiškai nustatoma, kiek būdingų taškų koordinatės pakinta po tam tikros emocijos pasireiškimo.

Veido išraiškų duomenų bazės

Darbe buvo naudotos dvi duomenų bazės: veidų bazė ColorFeret [4] (apie 1500 skirtingų žmonių) ir veido išraiškų bazė Cohn-Kanade AU-Coded Facial Expression Database [5] (apie 200 skirtingų žmonių veidų su maždaug 2000 išraiškų). Yra pateikti keli kiekvienos vieno asmens veido išraiškos vaizdo įrašai. Vaizdo įrašą sudaro nuo 10 iki 25 kadrų seka, kuri atspindi išraiškos kitimą nuo neutralios iki nusistovėjusios tiriamos išraiškos. Duomenų bazės kadrų sekos yra anotuotos FACS [6] kodavimo

sistema. Tyrimui buvo naudotos 325 sekos, atspindinčios penkias išraiškas.

Būdingų taškų bazės sudarymas

Remiantis tyrimais [3] kitų autorių [1,7] parinkta 18 veido taškų, kurių padėtis labiausiai kinta, keičiantis išraiškoms. Parinktas toks taškų išsidėstymas: antakiai – 6 taškai, akių kontūrai – 8 taškai, lūpų kontūras – 4 taškai. Tikslios taškų vietos parodytos 1 paveiksle.

Turimos Cohn-Kanade bazės kadrai leido išsamiai tyrinėti 5 pagrindinių išraiškų taškų padėčių pokyčius (pikta, linksma, neutrali, liūdna, nustebusi išraiškas). Pasirinktos išraiškos, kurių rasta daugiau nei 20 įrašų, vienareikšmiškai vaizduojančių tiriamą išraišką (2 pav.). Išraiškos ir tiriamų įrašų skaičiai pateikti 1 lentelėje.



1 pav. Taškai, kurių padėtis kinta keičiantis veido išraiškai: 1–8 akių ribos; 9–14 antakių ribos; 15–18 lūpų ribos

1 lentelė. Tirtų išraiškų kadrų duomenų bazė ir kadrų skaičius

| Išraiškos kodas | Išraiška | Įrašų skaičius |
|-----------------|-----------|----------------|
| 1. | Pikta | 64 |
| 2. | Linksma | 97 |
| 3. | Neutrali | 72 |
| 4. | Liūdna | 30 |
| 5. | Nustebusi | 62 |
| Iš viso: | | 325 |

Tyrimui pasirinkti pirmas ir paskutinis kadrai, kurie vaizduoja neutralią ir nusistovėjusią išraišką. Rankomis pažymėti taškai, automatiškai buvo patikslinti. Sukurta duomenų bazė iš pirmo (p) ir paskutinio (s) kadro

kiekvieno taško (n) koordinatinių ($X_{n,p}$, $X_{n,s}$, $Y_{n,p}$, $Y_{n,s}$) bei išraiškos kodo, gauto apdorojant FACS kodą.

Taškų koordinatinių pokyčių statistinė analizė












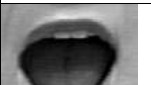
Sudaryta duomenų bazė apdorota statistiškai. Ieškoma būdingų taškų koordinatinių pokyčių, kurie būtų reikšmingi nustatant veido išraišką. Tirti kiekvieno taško santykiniai pokyčiai atskirai ir įvairių taškų porų tarpusavio pokyčiai.

Buvo apskaičiuotas santykinis kiekvieno taško neutralios (pirminės) ir tiriamos nusistovėjusios išraiškų koordinatinių pokyčiai:

$$\begin{cases} X_{i,m} = \frac{1}{n_i} \cdot \sum_{j \in D_n} \frac{x_{j,m}^{(s)} - x_{j,m}^{(p)}}{l_j}, \\ Y_{i,m} = \frac{1}{n_i} \cdot \sum_{j \in D_n} \frac{y_{j,m}^{(s)} - y_{j,m}^{(p)}}{l_j}. \end{cases} \quad (1)$$

čia n_i – i -osios išraiškos sekų skaičius; x, y – taškų koordinatės; m – taško indeksas $m=[1;18]$; X ir Y – vidutiniai santykiniai koordinatinių pokyčiai; l – kadre nekintantis atstumas tarp dviejų taškų (akių išorinių kampų tarp 1 ir 7 taško).

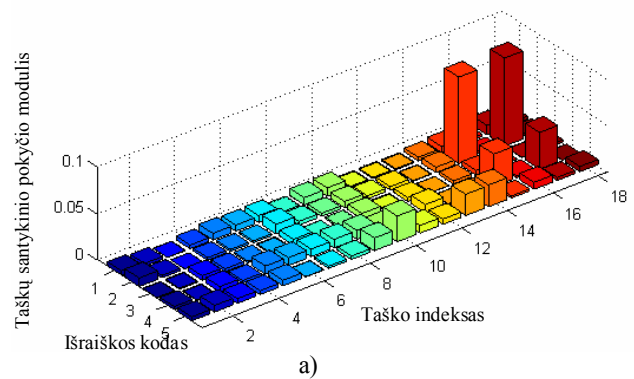
Buvo laikoma, kad taškų santykiniai pokyčiai pasiskirstę pagal normalinį dėsnį, todėl atliekant statistinius skaičiavimus taikyti tipiniai normalinio skirtingo dėsniai. Santykiai, nepatekę į $\pm 3\sigma$ intervalą, buvo atmetami.

| Išraiška | Akių zona | Lūpų zona |
|-----------------------|---|---|
| 0. Neutrali (pirminė) |  |  |
| 1. Pikta |  |  |
| 2. Linksma |  |  |
| 3. Neutrali |  |  |
| 4. Liūdna |  |  |
| 5. Nustebusi |  |  |

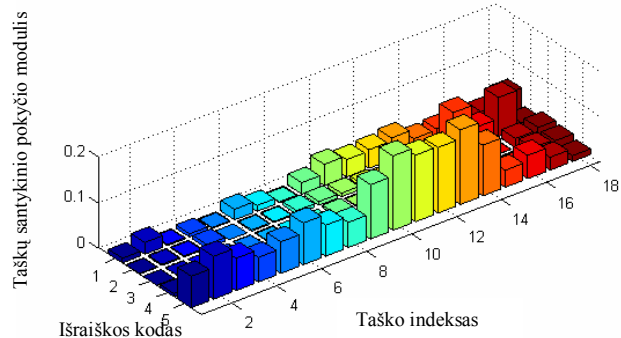
2 pav. Bazinės veido išraiškas rodantys kadrai

Atlikus skaičiavimus pagal (1) formulę, gauta, kiek vidutiniškai pakinta taško koordinatės nuo pirminės neutralios išraiškos iki nusistovėjusios. 3a ir 3b paveiksluose grafiškai parodyti atskirų išraiškų taškų santykiniai koordinatinių pokyčių moduliai.

Taškų reikšmingumui nustatyti buvo apskaičiuotos $X_{i,m}$ ir $Y_{i,m}$ dispersijos.



a)

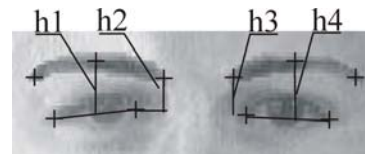


b)

3 pav. Santykiniai taškų pokyčiai, pereinant nuo pirminės į konkrečią galinę veido išraišką. Pateikiami taškų koordinatinių reikšmių moduliai: a) X ašimi; b) Y ašimi

Antrame etape vizualiai parinktos ir patikrintos taškų poros. Atstumų kitimas yra reikšmingas, nustatant išraišką. Tokias poras sudaro:

- atstumas vertikalia kryptimi tarp antakio vidinio taško ir akies vidinio taško (kairei akiai atstumo tarp 3 ir 11 projekcija į vertikaliąją ašį h2 (4 pav.), dešinei akiai – atstumo tarp 12 ir 5 taško projekcija į vertikaliąją ašį h3);
- „2–4“ ir „6–8“ akies plyšio aukštis, t. y. atstumas tarp apatinio ir viršutinio taško (kairei akiai – tarp 2 ir 4 taško (1 pav.), dešinei akiai – tarp 8 ir 13 taško);
- atstumas vertikalia kryptimi nuo antakių vidurio iki akies kampų koordinatinių vidurio h1 ir h4 (4 pav.).



4 pav. Taškų poros ir atstumai tarp jų, susiję su akimis

Kiti atstumai susiję su lūpomis:

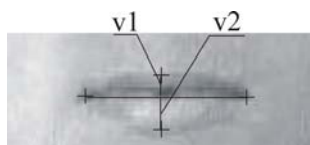
- „16–18“ vertikali linija tarp lūpos viršutinio (16) ir apatinio taško (18) (1 pav.);
- statinio aukštis nuo lūpos viršutinio taško (16) iki išorinius taškus jungiančios linijos (5 pav. linija v1);
- statinio aukštis nuo lūpos apatinio taško iki išorinius taškus jungiančios linijos (5 pav. linija v2);
- „15–17“ horizontalus atstumas tarp lūpos kampų (1 pav. tarp 15 ir 17 taško).

Santykinius neutralios ir nusistovėjusios išraiškų tarpusavio pokyčius apskaičiuojame pagal šias formules:

$$\delta v1 = \left(\frac{y_{16}^{(p)} - \frac{y_{15}^{(p)} + y_{17}^{(p)}}{2}}{y_{16}^{(s)} - \frac{y_{15}^{(s)} + y_{17}^{(s)}}{2}} - 1 \right) \cdot 100; \quad (2)$$

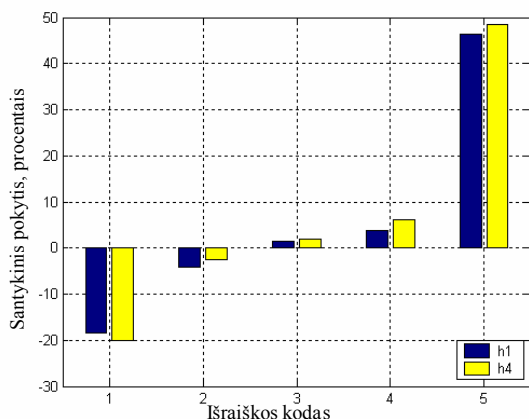
$$\delta v2 = \left(\frac{y_{18}^{(p)} - \frac{y_{15}^{(p)} + y_{17}^{(p)}}{2}}{y_{18}^{(s)} - \frac{y_{15}^{(s)} + y_{17}^{(s)}}{2}} - 1 \right) \cdot 100; \quad (3)$$

čia $y_{xx}^{(p)}$ – neutralios (pirminės) išraiškos xx taško y koordinatė, $y_{xx}^{(s)}$ – nusistovėjusios išraiškos xx taško y koordinatė.



5 pav. Lūpų taškų pokyčių atkarpos: v1 – viršutinės lūpos, v2 – apatinės lūpos

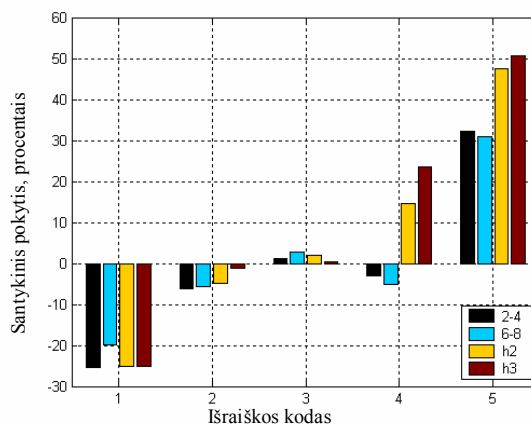
Antakių taškų judėjimas akių kampų atžvilgiu priklauso nuo konkrečios išraiškos. Labiausiai keičiasi nustebusi veido išraiška: antakis išsigaubia į viršų (6 pav. atkarpos h1 ir h4), susiraukusio žmogaus antakių vidinė dalis artėja prie akies kampo, trumpėja h2 ir h3 atkarpos (7 pav.).



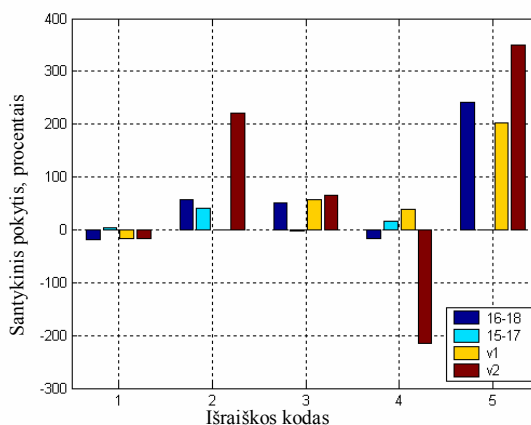
6 pav. Antakių centro vertikalūs pokyčiai h1 ir h4 esant įvairioms išraiškoms

Kitose išraiškose šių atkarpų kitimas nėra reikšmingas. Panašiai vertikalus akies plyšys reikšmingai keičiasi esant piktai ir nustebusiai išraiškoms (7 pav. „2–4“, „6–8“).

Lūpų formą apibūdinantis atstumas v2 reikšmingai keičiasi esant linksmai ir liūdnei išraiškoms. Nustebusiai išraiškai gerokai didėja atstumas tarp viršutinės (16) ir apatinės (18) lūpos (8 pav.). Horizontalus atstumas tarp lūpų kampų (15 ir 17 taškai) nežymiai kinta tik esant linksmai išraiškai – kampų taškai vienas nuo kito nutolsta.



7 pav. Akies plyšio pokytis „2–4“ ir „6–8“; antakio kampo ir akies kampo pokytis 2 ir h3 esant įvairioms išraiškoms



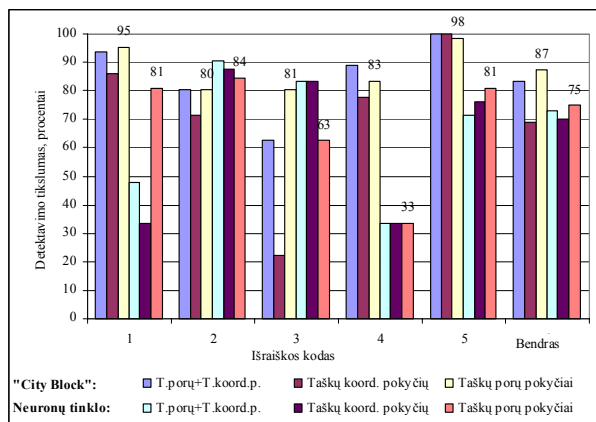
8 pav. Lūpų formos pokyčiai „16–18“; „15–17“; v1; v2 esant įvairioms išraiškoms

Gautų rezultatų vertinimas

Gauti rezultatai testuoti realiais duomenimis, sudarytais iš įvairių aprašytų išraiškų taškų. Išraiškos nustatomos dviem būdais: „City block“ [8] ir daugiasluoksnio perceptrono neuronų tinklu [9] naudojant aprašytus taškų koordinatinių poslinkių ir taškų tarpusavio pokyčių rezultatus. Ankstyvojo stabdymo (angl. early stopping) neuronų tinklas buvo apmokomas vienu, stabdomas antru ir tikrinamas trečiu duomenų rinkiniu. Abiem būdais gauti rezultatai buvo lyginami su tikru žinomu išraiškos rezultatu – jei algoritmo rezultatas sutapo su tikrąja išraiška, buvo laikoma, kad atpažinimas teisingas, jei nesutapo, – klaidingas.

Rezultatai rodo, kad geriausias bendras tikslumas 87 % „City block“ ir 75 % – neuronų tinklu (9 pav.) būdais gautas vertinant tik taškų porų pokyčius. Apskaičiuotas kiekvienos išraiškos detektavimo tikslumas rodo, kad skirtingos išraiškos detektuojamos nevienodai. 9 paveiksle pateikti kiekvienos išraiškos detektavimo rezultatai, kai naudojami skirtingi duomenų rinkiniai ir detektavimo būdai. Vertinant rezultatus, padaryta išvada, kad dėl atpažinimui naudojamų taškų koordinatinių santykinių

pokyčių iki 18 % padaugėja detektavimo klaidų. Neuronų tinklas sukelia didesnes atpažinimo klaidas, tačiau jos būna patikimesnės ir stabilesnės, nes mokoma ir testuojama skirtingais duomenų rinkiniais (naudojant tą patį rinkinį, gautas iki 98 % tikslumas, tačiau neuronų tinklas gali būti permokytas).



9 pav. Išraiškų detektavimo tikslumo rezultatai gauti, kai vertinamas taškų poros, taškų koordinacių pokytis ir kombinuotu būdu. 1–5 išraiškos kodas ir bendras (suminis) tikslumas

Išvados

Kiekvienas žmogus yra labai individualus ir emocijas rodo nevienodai. Vienų šypsėnos yra klasikinės: lūpų kampučiai pakyla ir lūpų linija išsitempia, kiti tai daro tik ištempdami lūpas. Todėl vienareikšmiškai nustatyti visas emocijas tik iš veido išraiškos yra sunku. Gaunama didelė duomenų sklaida ir svyravimas apie vidutinę vertę. Šiame darbe gauti rezultatai ir skaičiai prieš taikymą praktikoje turėtų būti tikslinami arba nustatomi detektuojant iš naujo.

Realių duomenų analizė rodo, kad „City block“ būdu gaunamas 87 % tikslumas, o neuronų tinklais – 75 %.

D. Dervinis. Automatinis veido išraiškos atpažinimas // Elektronika ir elektrotechnika. – Kaunas: Technologija, 2005.–Nr. 8(64). – P. 79–82.

Aprašytu būdu galima nustatyti veido išraišką statiniame kadre, identifikuoti vieną iš penkių tipinių žmogaus veido išraiškų (pikta, linksma, neutrali, liūdna, nustebusi). Naudojama aštuoniolika taškų, išdėstyti ant antakių, akių ir lūpų. Ištyrus 325 įvairių žmonių išraiškas, statistiškai nustatyta, kaip keičiasi šių taškų koordinatės keičiantis išraiškai nuo neutralios iki nusistovėjusios. Apskaičiuoti taškų porų tarpusavio pokyčiai, kurie yra reikšmingi keičiantis išraiškoms. Norint detektuoti išraišką aprašytu būdu, būtina panaudoti žmogaus neutralios veido išraiškos kadrus. Detektavimui buvo naudoti „City Block“ ir daugiasluoksnio perceptrono neuronų tinklo būdai. Pateiktų būdų bendras tikslumas siekia 87 %. Il. 9, bibl. 9 (lietuvių kalba; santraukos lietuvių, anglų, rusų k.).

D. Dervinis. Automatic Face Expression Detection // Electronics and Electrical Engineering. – Kaunas: Technologija, 2005.–No.8(64). – P. 79–82.

The method for detecting face expression in the static image is analyzed. Method is able to detect one from five typical face expressions (angry, happy, neutral, sadness, surprise). Algorithm uses eighteen points in eyebrows, eyes, lips. There were used 325 separate human expressions in experiment. Statistically were calculated changes of points coordinates in different expressions between neutrally and final expression images and defined important changes of points pairs. In order to detect face expression using this method, neutral expression must be taken in to account. In detection process were used „City Block“ and neural multilayer perceptron network methods. Total accuracy of the method is 87 %. Ill. 9, bibl. 9 (in Lithuanian; summaries in Lithuanian, English and Russian).

Д. Дервинис. Автоматическое распознавание выражения лица // Электроника и электротехника. – Каунас: Технология, 2005. – № 8(64). – С. 79–82.

Предложен способ распознавания выражения лица человека в статическом кадре. Этим способом идентифицируют одну из мин человека (злая, весёлая, нейтральная, печальная, удивлённая). Для распознавания использовано 18 точек. Указанным способом исследованы мины 325 человек, рассмотрено, как изменяются координаты указанных точек от нейтральной до рассматриваемой мины. Определены пары точек, изменения положения которых являются значительными для различных мин. Для детектирования использованы методы „City Block“ и многослойная перцептронная сеть. Точность предложенного метода достигает 87 %. Ил. 9, библи. 9 (на литовском языке; рефераты на литовском, английском и русском яз.).

Pastarasis detektavimo būdas yra kokybiškesnis ir stabilesnis. „City block“ naudoja vidutines gautas vertes ir nuokrypius nuo jų, neuroniniai tinklai apmokomi dideliais duomenų rinkiniais. Abu būdai tinkami taikyti praktikoje.

Ne visi pradžioje naudoti 18 taškų yra reikalingi. Kaip rodo 9 pav., nebūtina naudoti taškų koordinacių pokyčius, todėl galima atsisakyti akių 3, 5 ir antakių 9, 14 taškų.

Literatūra

1. **Bartlett M. S., Hager J.C., Ekman P.** Measuring facial expressions by computer image analysis // *Psychophysiology.* – Cambridge University Press, 1999.– No. 36.– P. 253–263.
2. **Liu Y., Schmidt K., Cohn J. F., Mitra S.** Facial asymmetry quantification for expression invariant human identification // *Computer Vision and Image Understanding,* 2003.–No. 91. – P. 138–159.
3. **Dervinis D.** Būdingų veido taškų išskyrimas dvimačiame vaizde // *Elektronika ir elektrotechnika.* – Kaunas: Technologija, 2004.–Nr. 5(54). – P. 16–20.
4. **The Color FERET Database.** Prieiga per internetą: <http://www.nist.gov/humanid/colorferet/colorferet.html>.
5. **Cohn–Kanade AU–Coded Facial Expression Database.** Prieiga per internetą: <http://www.cs.cmu.edu/~face>.
6. **FACS (Facial Action Coding System).** Prieiga per internetą: <http://face-and-emotion.com/dataface/>.
7. **Cohn J. F., Reed L., Moriyama T., Xiao J.** Multimodal coordination of facial action, head rotation, and eye motion during spontaneous smiles. In *Proceedings of the Sixth IEEE International Conference on Automatic Face and Gesture Recognition (FG'04).* – Seoul, Korea, 2004.
8. **Similarity and Distance.** [žiūrėta 2005–05–05]. Prieiga per internetą: <http://www.cs.bham.ac.uk/~slb/taxonomy/Taxonomy.html>
9. **Kohonen T.** *Self – organizing maps.* Berlin Springer, 2001. 161 p. – ISBN 3–540–67921–9.

Pateikta spaudai 2005 05 19

