

МОЖЛИВІСТЬ ВИКОРИСТАННЯ КОЛЬОРУ ПРИ РОЗПІЗНАВАННІ ЖЕСТІВ

Нюнькін К.М.

Донецький державний інститут штучного інтелекту

Пр. Богдана Хмельницького, 84, м. Донецьк, 83050

Тел.: (0622) 926082, факс: (0622) 926082

E-mail: kmn@iai.donetsk.ua

Розглядається проблема розпізнавання жестів у послідовності кольорових зображень. Зокрема, досліджується можливість використання кольору для визначення на зображеннях ділянок, що відповідають рукам. Показано обмеження використання для цих цілей кольору шкіри людини. Рекомендовано застосовувати для подачі команд жестами предмети насичених кольорів.

ВСТУП

Головна мета досліджень в галузі розпізнавання жестів полягає в тому, щоб створити систему, що може ідентифікувати певні людські жести і використовувати їх для передачі інформації або для керування різними пристроями.

Незважаючи на свою інтуїтивну зрозумілість, термін "розпізнавання жестів" різними дослідниками трактується по-різному. Дуже часто розпізнавання жестів ведеться з використанням спеціальних пристроїв, таких як рукавички з 10 - 20 датчиками (SensorGlove, PowerGlove, Cyberglove) [1, 2, 3] чи навіть за допомогою джойстика [4].

До області розпізнавання жестів відносять також стає останнім часом популярною задачу введення в комп'ютер з метою одночасного чи подальшого відображення формул, що малюються на звичайній дошці, схем, діаграм і т.п. [5]. Більш правильним представляється трактування поняття "розпізнавання жестів", яке полягає в тому, що приймання жестів людини здійснюється за допомогою відеокамери (аналогічно людському зору) і не вимагає застосування спеціальних пристроїв, що користувач повинен був би надягати або якими йому необхідно було б керувати.

Таким чином, при розпізнаванні жестів необхідно виявляти на зображенні руки людини і просліджувати зміну їхнього положення чи форми з часом. Для цього звичайно використовують або детектування руху [6, 7, 8], або виявлення областей, що мають колір шкіри, [9, 10] або обидва ці підходи використовують спільно [11].

Як правило жести, що виявляються, мають досить великий масштаб. Це можуть бути вказуючі жести, у яких задіяна вся рука, або рухи рук, подібні до крил, що махають [6]. У таких ситуаціях звичайно разом із кольором використовується детектування

областей руху.

При розпізнаванні дрібномасштабних жестів, у яких задіяні окремі пальці, кисть руки займає істотну частину кадру [7, 8]. Фон в цих випадках однорідний і контрастує по кольору зі шкірою людини.

1 ВИЯВЛЕННЯ НА ЗОБРАЖЕННІ ОБЛИЧЧЯ З ВИКОРИСТАННЯМ КОЛЬОРУ

Розглянемо критерії виділення на зображенні областей, що мають колір людської шкіри. Ця задача вирішується не тільки при розпізнаванні жестів, але і при виявленні обличчя людини. У той же час, дослідження з виявлення на кольорових зображеннях обличчя ведуться більш інтенсивно.

При виявленні на зображенні галузей, що мають колір шкіри, поряд зі звичайним RGB представленням кольору, тобто інтенсивностями червоної, зеленої і синьої складової кольору, використовується представлення, засноване на кольоровості - HSL представлення (Hue - колір або відтінок, Saturation - насиченість, Luminosity - яскравість):

$$H = \frac{\arctan(y/x)}{2\pi}, \quad S = \sqrt{x^2 + y^2}, \quad L = \frac{R + G + B}{3},$$

$$x = R - 0.5 \cdot (G + B), \quad y = \sqrt{3} \cdot (G - B).$$

Li & Kobatake [12] використовували колірний простір YIQ, що відповідає стандарту NTSC американського телебачення, для локалізації обличчя і його деталей. Для локалізації обличчя автори використовували закономірність, що "для жовтої раси в представленні YIQ обличчя має значення Q більше, а значення I менше, ніж темне оточення".

Як недоліки описаного підходу сегментації обличчя людини можна відзначити наступні моменти:

- У даному підході використані співвідношення між кольорами визначених елементів обличчя для конкретної "жовтої" раси. Для людей інших рас багато чого не має бути застосовано.
- Не враховується можлива наявність бороди, вусів, окулярів, нафарбованих губ і т.п.
- Фон і волосся покладаються темними, що не завжди справедливо.

Крім того, схоже, автори не праві, коли

стверджують, що "обличчя має значення Q більше, а значення I менше, ніж темне оточення". Насправді, як показують наші експерименти, закономірності виходять зворотні. Великі значення Q і маленькі I відповідають синім ділянкам, у той час як маленькі значення Q і великі I - жовтим.

Took & Craw [13] для виявлення обличчя водія в автомобілі, що займає істотну частину кадру, використовували колірний простір, аналогічний HSL. Формули для переведення RGB->HSL не приводяться, але вони відрізняються від загальноприйнятих, оскільки для кольору обличчя наводиться діапазон $150 \leq H \leq 200$ (звичайно H в області обличчя ~25-40), і, відповідно до приведених картинок, "колір" відрховується від зеленого, а не від червоного. Поряд із указаним діапазоном H, для області обличчя приймається наступний діапазон S: $128 \leq S \leq 224$, тобто зображення повинні мати добру колірну насиченість. За звичай значення S дуже рідко перевищують 100, а типові значення для області обличчя лежать у діапазоні 30-90.

Sirohey & Rosenfeld [14] для виявлення обличчя, що займає значну частину кольорового зображення з однорідним фоном, використовували дві нормалізовані компоненти RGB представлення:

$$r=R/(R+G+B)*100 \text{ і } g=G/(R+G+B)*100.$$

Автори провели дослідження, у яких на кольорових зображеннях вручну задавалася область обличчя і розраховувалися гистограми розподілу r і g. Отримана двовимірна гистограма апроксимувалася гаусіаном, у результаті чого для бази даних Aberdeen (UK), що містить 72 фотографії, було отримано, що центр розподілу відповідає точці $\mu=[42.91,32.28]$ з дисперсією $\sigma=[19.28,-5.45;-5.45,8.55]$.

Аналогічні дослідження були проведені для власної бази даних з 21 фотографії з не зовсім однорідним фоном. Для цієї бази даних були отримані трохи інші параметри гаусіана: $\mu=[52.66,29.99]$, $\sigma=[68.94,-30.89;-30.89,18.11]$.

Автори відзначають, що навіть незважаючи на норміровку, колір обличчя на фотографії залежить від умов освітлення, особливо від кольору джерела світла. У той же час вони вказують, що, якщо для даних умов освітлення визначити параметри гаусіана, те метод працює досить стійко.

Варто помітити, що параметри гаусіана залежать від умов освітлення, використовуваних камер і їхніх налаштувань. Тому визначити межі змін колірних компонентів (навіть нормалізованих) раз і назавжди не можна. Час від часу потрібна їхня корекція.

У технічному звіті лабораторії AT&T [15] для локалізації обличчя використовується три канали: 1) колір, 2) рух, 3) аналіз форми. Основна мета - скорочення часу при роботі з послідовністю кадрів. Звідси і використання каналу кольору як найшвидшого. Однак він також вважається і самим ненадійним. Тому його результати періодично підтверджуються аналізом руху.

Для колірного аналізу використовуються нормалізовані значення rgb, що дозволяє зменшити

залежність алгоритму від яскравості. При цьому темні пікселі, тобто ті, у яких сума $(R+G+B)<30$, обнуляються. Точні критерії вибору області обличчя не приводяться. Відповідно до малюнка розподілу відносних кольорів зображення в площині rg, на якому нахиленим еліпсом, що може бути уписаний у прямокутник $r=[125,\dots,175]$, $g=[50,\dots,90]$, обведена область, що відповідає "кольору шкіри".

Обличчя в даній роботі є домінуючим об'єктом на зображенні. Тому, якщо пікселів з кольором обличчя в одному зв'язному компоненту стане занадто мало чи занадто багато, такий аналіз повинен бути припинений.

Звернемо увагу на те, що автори не дуже довіряють колірної сегментації, постійно перевіряють її результати і мають критерії її припинення. Також варто врахувати, що колірна сегментація вимагає виконання попереднього калібрування колірного простору.

В авторефераті Талеба з ИТК НАН Беларусі [16] при ідентифікації ділянок шкіри людини на кольорових зображеннях також використовувався колірний простір HSL разом з деякими комплексами, складеними з компонентів R, G і B.

Використаний ним набір ознак, з одного боку, виглядає трохи перевантаженим, а з іншого боку - не містить насиченість S.

Аналізуючи роботи з виявлення обличчя людини на основі інформації про колір, не можна не помітити наступні моменти:

1. Колір шкіри в різних людей (навіть однієї раси) варіюється в досить широких межах.
2. Колір шкіри залежить від умов освітлення.
3. Перед виконанням виявлення обличчя цим методом бажано виконати калібрування колірного простору.
4. Найчастіше при виявленні обличчя канал кольору використовується разом з іншими: рухом, градієнтом яскравості, формою, аналізом оточення.

Імовірно, зазначені обмеження і послужили причиною того, що такі методи не знайшли широкого застосування.

2 ВИЯВЛЕННЯ НА ЗОБРАЖЕННІ РУК ЛЮДИНИ

Для дослідження колірних характеристик областей обличчя і рук людини була розроблена спеціальна комп'ютерна програма, що дозволяє вводити кольорові зображення, одержувані підключеної до комп'ютера веб-камери, а також із вже існуючих файлів. Програма розраховувала компоненти різних представлень кольору (RGB, HSL, YIQ) і дозволяла виділяти на зображенні ділянки, що відповідають різним критеріям, складеним з цих компонентів.

Експерименти, проведені з використанням указаної програми, показали, що

- Значення кольору (H) шкіри людини можуть лежати в діапазоні від 4 до 40.
- Значення насиченості (S) шкіри людини можуть лежати в діапазоні від 30 до 90.
- Колір шкіри у різних людей варіюється в досить широких межах.
- Колір шкіри навіть однієї людини залежить від умов освітлення.
- Одержуваний на зображенні колір шкіри людини залежить від налаштувань камери, що найчастіше виконуються автоматично і не контролюються.
- Колір, близький до кольору шкіри людини, можуть мати деталі одягу, що оточують предмети і стіни приміщення (див. таблицю 1).

Таблиця 1 - Об'єкти, що мають колір, близький до кольору шкіри людини (H=4-40)

Об'єкт	Значення H	Об'єкт	Значення H
Стіни	17-22	Дошки	7-27
Штори	17-19	Бетонні доріжки	20-30
Цегельний будинок	5-18	Камені	17-22
Автобус	14-20	Пісок	20-25
Підфарники	9-20	Листи	31-45

Таким чином, використання кольору шкіри людини при створенні системи розпізнавання жестів для роботи в реальних умовах (неконтрольовані умови освітлення, різні люди оператори, довільний навколишній фон) виглядає дуже проблематичним. Урахування руху як додаткового критерію при виявленні рук людини мало що змінює, оскільки в реальних умовах у кадрі можуть рухатися не тільки

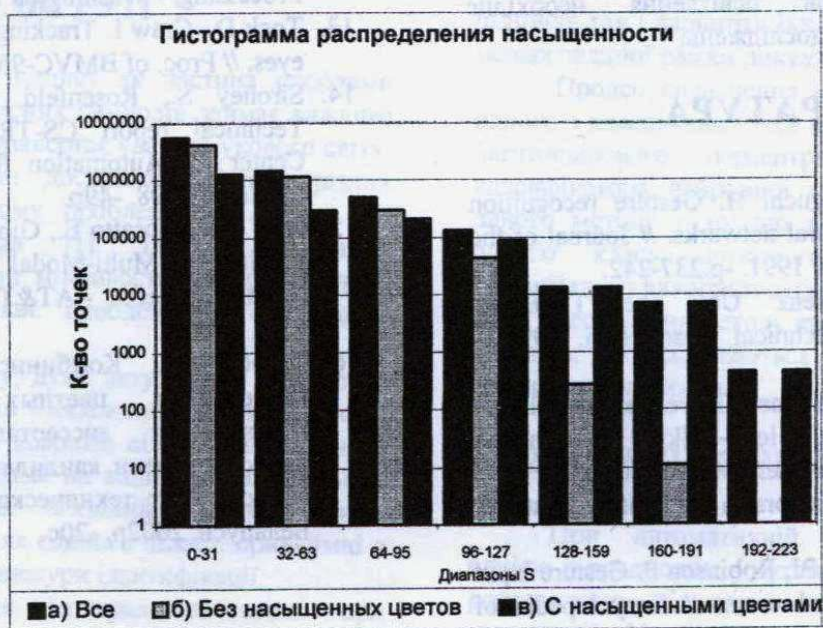
руки оператора, але і навколишні об'єкти.

У той же час, у проведених експериментах, у тому числі і поза приміщенням, було відзначено вкрай мало об'єктів, що мають насичені кольори (зі значеннями S, що перевищують середину можливого діапазону, тобто 128). Це були деталі одягу і яскраві квіти, у яких значення S перевищували 160. А от колір зелені в переважній більшості випадків не дуже насичений: значення S найчастіше не перевершує 70, хоча у яскраво освітленої сонцем свіжої зелені може досягати 150 (з 255 можливих). У таблиці 2 наведені об'єкти, які мають високі значення насиченості, що виявляються поза приміщенням у ясний сонячний день. Слід зазначити, що такі об'єкти були виявлені тільки на 19 кадрах з 71, хоча оператор знімав саме ті об'єкти, що на його думку мали яскраві насичені кольори.

Таблиця 2 - Об'єкти, що мають високі значення насиченості (S>128)

Об'єкт	Значення S	Значення H
Червона машина	150-190	250-256
Жовтий автобус	130-160	17-24
Свіже листя на деревах	130-150	49-57
Червоні троянди	160-210	245-256
Зелена вода в басейні	130-135	42-46
Рожева блузка	140-150	237-245
Малинова сукня	130-160	235-247
Червона блузка	130-180	250-255

На мал. 1 наведені гістограми розподілу насиченості на зображеннях, знятих поза приміщенням у ясний сонячний день. Можна помітити, що кількість точок, що мають дану насиченість, зменшується з її ростом за показовим законом.



Малюнок 1 - Гістограми розподілу насиченості:

- а) усі зображення, б) зображення, що не містять насичених кольорів, в) зображення з об'єктами, що мають насичені кольори.

Наведені вище результати експериментів говорять про те, що при виявленні рук оператора можна використовувати насичені кольори, для чого варто застосовувати надягнуті на руки рукавички визначеного насиченого кольору. Поріг насиченості при цьому, імовірно, буде залежати від яскравості освітлення.

ВИСНОВКИ

Використання кольору шкіри людини при виявленні на зображенні обличчя і рук оператора має певні істотні недоліки, головними з яких можна вважати різниці кольору цих об'єктів у різних людей, залежність кольору від умов освітлення, наявність великої кількості інших об'єктів, що мають схожі кольори. І якщо перші два можуть бути усунуті калібруванням колірного простору, проведеної перед кожним сеансом роботи, то останній у випадку неможливості усунення з кадру сторонніх об'єктів, близьких по кольору до шкіри людини, перетворюється в серйозну проблему. Та й постійне виконання калібрування теж не завжди можливо. У результаті використання кольору при виявленні на зображенні обличчя і рук оператора має обмежену область застосування і найчастіше використовується разом з іншими критеріями: рухом, градієнтом яскравості, формою, аналізом оточення.

З іншого боку, проведені експерименти по дослідженню колірних характеристик різних об'єктів показали, що вкрай мало об'єктів мають насичені кольори. При виявленні рук оператора можна використовувати цей факт, для чого варто застосовувати надягнуті на руки рукавички визначеного насиченого кольору. Для рішення питання про поріг насиченості і його можливий залежності від умов освітлення необхідне проведення додаткових досліджень.

ЛІТЕРАТУРА

1. Murakami K., Taguchi H. Gesture recognition using recurrent neural networks. // Journal of the ACM, 1(1), January 1991. -p.237-242.
2. Hofmann F., Henz Ue. The TU-Berlin SensorGlove. - Technical Description. 1995. - http://pdv.cs.tu-berlin.de/forschung/SensorGlove2_engl.html.
3. Fels S., Hinton G. Glove-TalkII: An Adaptive Gesture-to-Formant Interface. // Proc. of the 1995 Conf. on Human Factors in Computing Systems, May 1995.
4. Keates S, Clarkson PJ, Robinson P. Gesture input and motion-impaired users. // Encyclopedia of Library and Information Science, Marcel Dekker, Inc., 2001, Vol.70. -p.128-150.
5. Damm, C.H., Hansen, K.M., Thomsen, M. Tool Support for Object-Oriented Cooperative Design:

Gesture Based Modeling on an Electronic Whiteboard. // Proc. of CHI'2000. ACM Press, 2000.

6. Cutler R., Turk M. View-based Interpretation of Real-time Optical Flow for Gesture Recognition. // IEEE Int.Conf.on Automatic Face and Gesture Recognition, April 1998, Nara, Japan.
7. Laptev I., Lindeberg T. Tracking of multi-state hand models using particle filtering and a hierarchy of multi-scale image features.- Technical report CVAP245, ISRN KTH NA/P--00/12--SE. Department of Numerical Analysis and Computer Science, KTH (Royal Institute of Technology), Stockholm, Sweden, September 2000.
8. Gupta D. Computer Gesture Recognition: Using the Constellation Method. // Caltech Undergraduate Research Journal, 2001, vol.1, N1. - pp.26-31.
9. Marcel S., Bernier O., Viallet J.-E., Collobert D. Hand Gesture Recognition Using Input-Output Hidden Markov Models. // Proc. of the Fourth IEEE Int. Conf. on Automatic Face and Gesture Recognition 2000.
10. Bretzner L., LaptevI., Lindeberg T. Hand Gesture Recognition using Multi-Scale Colour Features, Hierarchical Models and Particle Filtering. // Proc. of the 5th Int. Conf. on Automatic Face and Gesture Recognition, 2002.
11. McKenna S., Gong S. Gesture Recognition for Visually Mediated Interaction using Probabilistic Event Trajectories. // Proc. of BMVC'98, 14-17 September 1998, Southampton, England.
12. Li Y., Kobatake H. Extraction of Facial Sketch Images and Expression Transformation Based on FACS. // Proc. of IEEE Int. Conf. on Image Processing. - Washington D.C., 1995, v3.
13. Tock D., Craw I. Tracking and measuring drivers eyes. // Proc. of BMVC-95, -p.383-392.
14. Sirohey S., Rosenfeld A. Eye Detection. - Technical report CS-TR-3971 (CAR-TR-896) Center for Automation Research University of Maryland, 1998. -49p.
15. Graf H.P., Cosatto E., Gibbon D., Kocheisen M., Petajan E. Multi-Modal System for Locating Heads and Faces. - AT&T Lab Technical Report 95.5.1, 1996.
16. Талеб М.А. Комбинированные алгоритмы сегментации цветных изображений. - Автореферат диссертации на соискание ученой степени кандидата технических наук. Минск, Ин-т технической кибернетики НАН Беларуси. 2002р. -20с.