

УДК 007:61, 681.3.01

М.Ю. Катаев, Н.В. Ковалев

Оценка положения головы человека по анализу изображений

Рассматривается применение программно-аппаратного комплекса для задач исследования особенностей двигательной активности головы человека по набору изображений для медицинских целей. Описаны алгоритмы выделения лица человека и соответственно положения головы человека, а также тестовые результаты обработки изображений.

Ключевые слова: изображение, методы обработки, положение головы человека.

Введение

Голова человека является характеристикой, по которой можно узнать как психологические, так и физиологические особенности [1–3]. Древние врачи использовали наблюдение особенностей лица и головы как некие индикаторы состояния человека. Однако эти знания ушли из общей практики врачей и остались лишь в небольших группах, не связанных с официальной медициной. В настоящее время известно несколько приложений, где используется изображение головы человека, в том числе и его лица:

- 1) исследовательская работа (например, проводимая в <http://graphicon.ru>);
- 2) оперативно-розыскная деятельность (фотографии, видеосъемка);
- 3) наблюдение, видеоконтроль (фотографии, видеосъемка);
- 4) Интернет поисковые машины (фотографии, например по <http://image.google.com>);
- 5) медицина (фотографии и видеозображения).

Однако во всех этих приложениях, кроме первого, важен лишь процесс общего опознания лица человека или формы головы (в виде эллипса, некоторой маски и др.), без однозначного выделения каких-либо его особенностей. В данной статье рассматривается решение, связанное с детектированием лица человека и последующим определением характеристик перемещения центра тяжести головы человека. Лицо человека детектируется по нескольким параметрам: цвет кожи и характерные точки лица (глаза, нос и рот).

Определение особенностей лица связано с решением задач: предобработка (фильтрация шума и цветовая коррекция), обнаружение лица на изображении (нахождение контура или сегментация) и последующее выделение лица (удаление фона). Человек достаточно просто определяет в типичных условиях освещения и расстояний задачу обнаружения лица, но при построении автоматизированной системы обнаружения лица возникают сложности, которые необходимо учитывать в алгоритмах: 1) сильно отличные антропологические характеристики лица у разных людей; 2) изменение ориентации лица относительно камеры; 3) возможное присутствие индивидуальных особенностей (усы, борода, очки, брови различной густоты и т.д.); 4) выражение лица; 5) условия съемки (освещение, качество камеры). Все эти и некоторые другие проблемы в значительной степени влияют на изображение лица, а значит, и на процесс его определения на изображении.

Данная статья представляет результаты, полученные на первой стадии решения поставленной задачи, а именно: использование вебкамеры как источника изображений, выделение из видеопотока статических изображений с форматом BMP и последующая их предобработка, обработка с целью получения двигательной активности головы человека.

Описание аппаратно-программного комплекса

Программно-аппаратный комплекс (ПАК) состоит из вебкамеры (стандартное разрешение 640×480 с частотой 25–30 кадров в секунду) и программных модулей различного назначения (рис. 1). Первый модуль «StaticImage» предназначен для выделения из видеопотока статических изображений (типа BMP) через заданные промежутки времени. Промежутки времени зависят от задачи исследований (антропометрические измерения, динамика участков лица или др.). Этот модуль позволяет считывать видеофайлы (в формате AVI) или изображения (в формате BMP), полученные ранее. Далее набор изображений синхронизируется по яркостной картине с помощью методики «серый мир» и других преобразований в блоке «ImagePreProc». Полученный набор изображений сжимается, записывается в базу данных изображений «ImageDB». Следующий модуль «FaceColor» служит для выделения области головы (лица) человека и удаления фона (придание фону белого цвета). Модуль «HeadMove» предназначен для оценки двигательной активности головы, т.е. оценки времени и положения головы во время получения изображений или видеофайла. Модуль «Анализ» предназначен для анализа состояния лица и положения головы согласно поставленной задаче исследования.

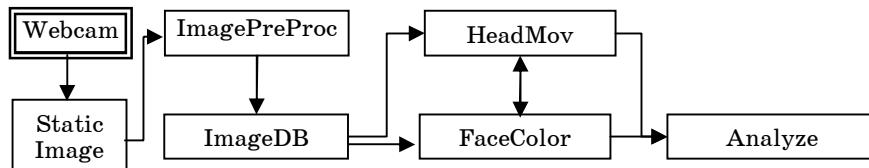


Рис. 1. Блок-схема программно-аппаратного комплекса

Применяемые алгоритмы

На первом шаге включается методика «серый мир», согласно которой сумма всех цветов на изображении сцены дает серый цвет. Для этого необходимо рассчитать средние яркости по всем каналам [1]:

$$\begin{aligned} \bar{R} &= \frac{1}{Nx \cdot Ny} \sum_{x,y} R(x,y); \quad \bar{G} = \frac{1}{Nx \cdot Ny} \sum_{x,y} G(x,y); \quad \bar{B} = \frac{1}{Nx \cdot Ny} \sum_{x,y} B(x,y); \\ W &= \frac{\bar{R} + \bar{G} + \bar{B}}{3}; \\ R' &= \frac{R \cdot W}{\bar{R}}; \quad G' = \frac{G \cdot W}{\bar{G}}; \quad B' = \frac{B \cdot W}{\bar{B}}; \end{aligned}$$

здесь $R(x,y), G(x,y), B(x,y)$ – цветовые характеристики изображения (красный, зеленый и синий); Nx, Ny – размер изображения в пикселях.

После этой процедуры необходимо преобразовать цветовую палитру согласно указанным ниже формулам, что позволяет выровнять контрастность изображений [4]:

$$R'' = \frac{(R' - R_{\min}) \cdot 255}{(R_{\max} - R_{\min})}; \quad G'' = \frac{(G' - G_{\min}) \cdot 255}{(G_{\max} - G_{\min})}; \quad B'' = \frac{(B' - B_{\min}) \cdot 255}{(B_{\max} - B_{\min})},$$

где $(R, G, B)_{\max}, (R, G, B)_{\min}$ – максимальная и минимальная величина цвета на изображении из диапазона [0–255].

Далее идет следующий этап предобработки, конвертация цветовой палитры RGB в YIQ, или в обратном порядке, по формулам [1]:

$$\begin{bmatrix} Y \\ I \\ Q \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ 0.596 & -0.274 & -0.322 \\ 0.211 & -0.522 & 0.311 \end{bmatrix} \begin{bmatrix} R'' \\ G'' \\ B'' \end{bmatrix}; \quad \begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} 1 & 0.956 & 0.623 \\ 1 & -0.272 & -0.648 \\ 1 & -1.105 & 0.705 \end{bmatrix} \begin{bmatrix} Y \\ I \\ Q \end{bmatrix}.$$

Фильтрация шума проводится с помощью свертки с функцией Гаусса [4]:

$$Y'(i,j) = \sum_{l=-n}^n \sum_{k=-m}^m Y(i+l, j+k) \cdot G(l, k, \sigma); \quad G(l, k, \sigma) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{d^2}{2\sigma^2}}, \quad d = \sqrt{l^2 + k^2},$$

здесь n, m – размер окна сглаживания (типовично $m=n=3\div7$); σ – ширина размытия, выражаемая в пикселях и связанная с размером окна.

Существующие алгоритмы выделения лиц на изображении можно разделить на три категории. К первой относятся методы, использующие цветовые характеристики кожи человека, ко второй категории можно отнести методы с использованием формы (это эллипс, перекрещенные отрезки (глаза–нос и/или рот) или прямоугольник или др.) и к третьей – методы на основе сравнения с образом, с использованием метода главных компонент, сингулярных разложений и др. [1–2]. Нами ввиду благоприятных условий измерений, которые легко реализуются в медицинских учреждениях, использовано два первых критерия. За основу первого критерия был взят алгоритм поиска цвета кожи человека [1] по формуле

$$\begin{aligned} \left(\frac{G}{R} \leq 0,839 \right) \text{ и } \left(\frac{G-B}{R+G+B} \leq 0,054 \right) \text{ и } \left(\frac{G \cdot B}{(R+G+B)^2} > 0,067 \right) \text{ и} \\ \left(\frac{G \cdot B}{(R+G+B)^2} < 0,098 \right) \text{ и } \left(\frac{B}{G} \leq 1,048 \right) \text{ и } \left(\frac{G}{3(R+G+B)} \leq 0,108 \right). \end{aligned}$$

Согласно указанным коэффициентам изменения цветов пикселей на изображении происходит сегментация, где все пиксели, которые соответствуют условию, индексируются как «1», а не соответствующие – «0».

После того как проведено разбиение изображения на области, происходит выделение контура изображения, чтобы разделить изображение на две части: лицо и фон. Нами используется тот факт, что резкий перепад яркости представляет собой границу фон-лицо (при условии однородного фона). На лице человека все его особенности имеют различные оттенки, что также приводит к возможности определять градиентными методами наличие бровей, глаз, носа и т.д.

Лицо человека имеет следующее описание, особенности каждого элемента которого представляют интерес для медиков и соответственно должны быть представлены в разрабатываемом ПАК. Исследуется лицо (форма, вертикальный и горизонтальный профиль, полнота, соотношение высоты и ширины), лоб, нос (высота, ширина, выступание, переносы, спинка носа, основание носа, кончик носа, крылья носа, ноздри), брови, глаза (форма, цвет, отстояние от переносицы), веки, рот (размер, контур, положение углов), скулы, подбородок, ушные раковины. Однако все эти параметры достаточно сложно автоматически определить из набора изображений с помощью известных математических алгоритмов.

Для решаемой нами задачи достаточно выделить лишь лицо на изображении и контролировать его перемещение по кадру. Временные и пространственные особенности перемещения и представляют основной интерес.

В работе использован двухуровневый подход, в котором сначала локализуется лицо по вышеописанному алгоритму (цвет кожи) и затем производится уточнение. На втором этапе для локализации лица на изображении применяется метод проекций [4–6]. Одновременно этот подход позволяет определить местоположение таких основных элементов лица как глаза, нос и рот. Так как изображение бинаризовано, на месте особенностей лица наблюдаются сгущения точек, а значит, в гистограмме наблюдаются максимумы. Он заключается в том, что рассчитываются 2 функции, значениями одной являются интегралы яркости изображения по вертикали (H), а второй – по горизонтали (V), в результате получаем векторы:

$$H(x) = \sum_{y=y1}^{y1+N} I(x, y), \quad V(y) = \sum_{x=x1}^{x1+M} I(x, y),$$

где N, M – количество точек изображения.

Соответственно, находятся максимумы гистограммы и сравнивается их положение со стандартными характеристиками, например лицо делится на две части по вертикали, как расстояние между глазами, размер носа также примерно равен этому состоянию и т.д.

Пример применения данного подхода представлен на рис. 2. Сплошными линиями показаны положения максимумов и минимумов гистограмм.



Рис. 2. Пример вычисления гистограмм и бинаризации изображения

Описание программы

Для реализации вышеописанных алгоритмов нами разработана программа, которая базируется на блоках, представленных на рис. 1. В настоящей версии программы реализованы блоки чтения видеопотока, выделения изображений через заданные промежутки времени, записи их в базу данных, предобработки, выделения фона и анализа. Реализована возможность чтения сторонних видеозаписей в формате AVI. Разработка базы знаний (медицинских) продолжается, является длительной и требует особой внимательности и тщательности, так как связана со здоровьем человека и поэтому не представлена в данной части статьи. Также пока оценка состояния и анализ являются незавершенными элементами программы, находящимися в стадии разработки, что связано с сопоставлением знаний математических и медицинских. Результаты подобных исследований будут приведены во второй части статьи.

Для тестовых расчетов нами использовалась открытая база изображений HPIID [7] (Head Pose Image Database), в которой содержится 15 наборов, состоящих из 93 полуточновых изображений разных людей. В базе даны изображения (пример см. на рис. 3), в формате jpg для двух направлений:

Вертикальный угол = $\{-90, -60, -30, -15, 0, +15, +30, +60, +90\}$;

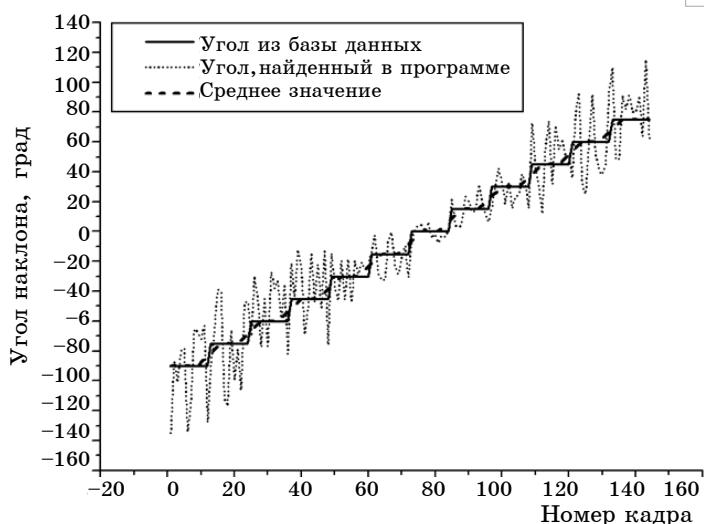
Горизонтальный угол = $\{-90, -75, -60, -45, -30, -15, 0, +15, +30, +45, +60, +75, +90\}$.



Рис. 3. Серия тестовых изображений поворота головы человека на заданный угол

Один из элементов модуля «Анализ» предназначен для анализа состояния лица согласно задаче исследования. На данный момент модуль позволяет выделить глаза и нос человека и по соотношениям сторон получившегося треугольника определить углы поворота и наклона головы. Проведенные тесты позволяют утверждать, что описываемый программный комплекс дает возможность с удовлетворительной точностью (в среднем ошибка 1–3 град) определять углы наклона вперед-назад и поворота головы в следующих диапазонах:

- наклоны головы: от -30 до 35 град,
- поворот головы: от -15 до 15 град.

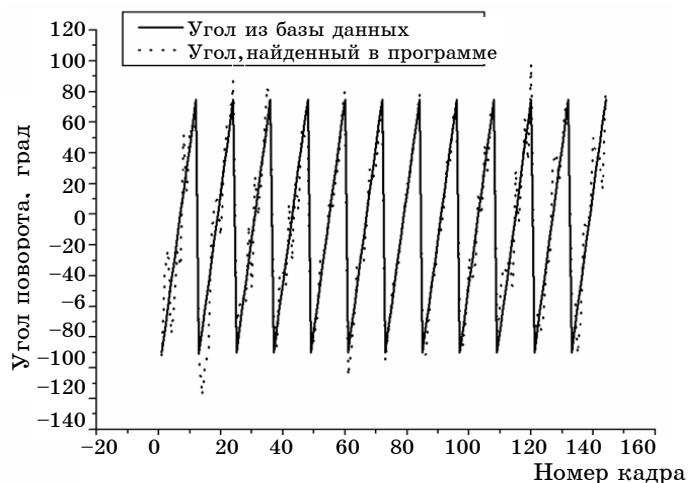


Результаты работы программы приведены на рис. 4 и 5. На рис. 4 показано сравнение рассчитанных углов наклона головы, а на рис. 5 – сравнение рассчитанных углов поворота головы в программе с тестовыми значениями, приведенными в базе данных.

Рис. 4. Сравнение рассчитанных углов наклона головы в программе с тестовыми значениями, приведенными в базе данных [7]

Естественно на результат работы могут влиять такие особенности человеческого лица, как прическа (например, длинная челка), которые не позволяют точно определить расположение некоторых характерных точек лица. Эти и другие особенности обработки изображений планируется решать в ходе развития данной тематики.

Рис. 5. Сравнение рассчитанных углов поворота головы в программе с тестовыми значениями, приведенными в базе данных [7]



Заключение

В статье представлен разработанный программно-аппаратный комплекс, позволяющий на основе обработки потока изображений от вебкамеры проводить оценку углов поворота и наклона головы человека. Получены первые тестовые результаты, которые подтверждают работоспособность предлагаемых способов оценки искомых параметров. Найдено, что повороты головы оцениваются более точно (средняя погрешность 1–3 град), чем наклоны головы (средняя точность 3–5 град).

Выражаем благодарность зав. неврологической клиникой СИБГМУ Н.Г. Катаевой и зав. неврологическим стационаром Томской больницы ФГУ «СОМЦ Росздрава» В.А. Чистяковой за консультации в ходе работы над проектом.

Литература

1. Sigal L. Skin color-based video segmentation under time varying illumination / L. Sigal, S. Sclaroff, V. Athitsos // IEEE Transactions on Pattern Analysis and Machine Intelligence. – 2004. – Vol. 28, N. 7. – P. 862–877.
2. Пальчун В.Т. Тест высокочастотных колебаний головы в исследовании вестибуло-окулярного рефлекса у здоровых лиц / В.Т. Пальчун, С.Н. Деревянко // Вестник оториноларингологии. – 2000. – № 2. – С. 4–8.
3. Eysenck H.I. Personality structure and measurement / H.I. Eysenck, S.R. Eysenck. – London.: Routledge and Kegan, 1969. – 300 р.
4. Прэтт У. Цифровая обработка изображений: В 2 кн. / Под ред. Д.С. Лебедева. – М.: Мир, 1982. – 620 с.
5. Сойфер В.А. Методы и алгоритмы Ч. 1. // Компьютерная обработка изображений. – СОЖ. – 1996. – № 2. – С. 118–124.
6. Сойфер В.А. Методы и алгоритмы Ч. 2. // Компьютерная обработка изображений. – СОЖ. – 1996. – № 3. – С. 110–121.
7. Gourier N. Estimating face orientation from robust detection of salient facial features / N. Gourier, D. Hall, J.L. Crowley // Proceedings of Pointing 2004, ICPR, International Workshop on Visual Observation of Deictic Gestures. – Cambridge, UK, 2004. – P. 124–128.

Катаев Михаил Юрьевич

Д-р техн. наук, профессор каф. автоматизированные системы управления ТУСУРа

Рт. тел.: (382-2) 70-15-36

Эл. почта: kmy@asu.tusur.ru

Ковалев Никита Викторович

Аспирант каф. автоматизированных систем управления ТУСУРа

Kataev M.Yu., Kovalev N.V.

Application of video-images in studying the human movement activity

The article presents the description of a developed hardware-software complex for the research of the movement activity of human head, for medical use. The techniques used for image processing are shown in the article.

Keywords: image, processing methods, human head position.