

УДК 004.932

АЛГОРИТМ РАСПОЗНАВАНИЯ ЖЕСТОВ РУК НА ОСНОВЕ СКЕЛЕТНОЙ МОДЕЛИ КИСТИ РУКИ

А. В. Носов

Сибирский государственный аэрокосмический университет имени академика М. Ф. Решетнева
Российская Федерация, 660014, г. Красноярск, просп. им. газеты «Красноярский рабочий», 31
E-mail: alexander@nosov.org

Рассматривается задача распознавания жестов кисти руки по статическому изображению. Предложен метод сегментации кисти руки, основанный на цветовой модели кожи в цветовом пространстве YCrCb. Введены понятия «непрерывный скелет многоугольной фигуры», «радиальная функция скелета» и «максимальный пустой круг». Предложен метод распознавания жестов, основанный на анализе формы непрерывного скелета кисти руки и выделении признаков, классифицирующих образ жеста. Введено понятие центра ладони и предложен метод ее вычисления, основанный на непрерывном скелете ладони. Предложен метод вычисления координат начала и кончиков пальцев на основе непрерывного скелета и значения его радиальной функции.

Ключевые слова: обработка изображений, распознавание жестов рук, сегментация изображений, непрерывный скелет.

ALGORITHM OF HAND GESTURE RECOGNITION BASED ON SKELETON MODEL OF HAND

A. V. Nosov

Siberian State Aerospace University named after academician M. F. Reshetnev
31, Krasnoyarsky Rabochy Av., Krasnoyarsk, 660014, Russian Federation
E-mail: alexander@nosov.org

The task of hand gesture recognition in the context of image is considered in this article. Hand segmentation method, based on analysis color skin in the YCrCb color model is used. Following terms are defined: "Topological skeleton of polygonal figure", "radial function of topological skeleton", "maximum empty circle". Gestures recognition method based on the topological skeleton analysis and extraction essential for recognition features is offered. Center palm term is introduced and its calculation method based on the topological skeleton is offered. Moreover this article describes fingers origin and finger-tips coordinates calculation method based on the topological skeleton and its radial function value.

Keywords: image processing, hand gesture recognition, image segmentation, topological skeletons.

В области машинного зрения одной из перспективных задач является задача распознавания жестов рук. О важности данной задачи говорит большое количество работ в данной области [1–5]. Практическое применение технологии распознавания жестов рук включает в себя приложения виртуальной реальности, распознавания языка глухонемых, человеко-машинного взаимодействия и др. Задачу распознавания жестов можно разделить на два этапа: локализация кисти руки и распознавание образа жеста. Оба этапа являются сложными и трудоемкими задачами.

В большинстве случаев жесты классифицируют на два типа: статические и динамические. Под статическим жестом понимается состояние ладони в определенный момент времени. Статические жесты иногда называют позами, их распознавание происходит на статическом изображении. Динамические жесты являются последовательностью статических жестов, и

распознавание их происходит на видеопоследовательности.

В данной статье рассматривается задача распознавания жестов кисти руки по статическому изображению. Для решения поставленной задачи в статье предлагается использовать метод сегментации, основанный на цветовой модели кожи, и метод распознавания, основанный на анализе формы кисти руки и выделении признаков, классифицирующих образ жеста. Анализ формы выполняется на основе непрерывного скелета. Скелет ладони позволяет анализировать направление жеста в пространстве, определить состояние пальцев (разжаты или сжаты в кулак). Инвариантные признаки являются дескрипторами формы ладони, позволяющими в быстрой степени производить поиск схожих жестов по имеющейся базе данных эталонных жестов. Предлагаемый алгоритм работает с силуэтом кисти руки, не учитывая текстуру ладони,

что позволяет распознавать жесты людей различных рас и работать с изображениями низкого качества, полученными с веб-камер или в условиях слабой освещенности.

Сегментация по цвету кожи [6–8] представляет собой преобразование RGB -изображения в другое цветное пространство, в котором различные цвета кожи расположены близко друг к другу. Чем ближе цвета кожи лежат друг к другу, тем качественнее будет проведена сегментация. Лучшим для сегментации кожи является пространство $YCrCb$ [6]. Переход в пространство $YCrCb$ из пространства RGB осуществляется следующим способом:

$$\begin{pmatrix} Y \\ Cr \\ Cb \end{pmatrix} = \begin{pmatrix} 0,299 & 0,587 & 0,114 \\ -0,169 & -0,331 & 0,500 \\ 0,500 & -0,419 & -0,081 \end{pmatrix} \cdot \begin{pmatrix} R \\ G \\ B \end{pmatrix}, \quad (1)$$

где Y – яркостная компонента; Cr и Cb – хроматические компоненты.

Уравнения линий, определяющих границы цвета кожи, представлены ниже:

$$\left\{ \begin{array}{l} Cr \geq -2(Cb + 24); \\ Cr \geq -4(Cb + 32); \\ Cr \geq -(Cb + 17); \\ Cr \geq 25(Cb + Q_1); \\ Cr \geq Q_3; \end{array} \right. \left\{ \begin{array}{l} Cr \leq \frac{220 - Cb}{6}; \\ Cr \leq \frac{4}{3}(Q_2 - Cb); \\ Cr \geq 0,5(Q_4 - Cb), \end{array} \right. \quad (2)$$

где Q_1, Q_2, Q_3 и Q_4 являются вспомогательными переменными, которые вычисляются по формулам

$$\left\{ \begin{array}{l} Y > 128, \\ Y \leq 128, \end{array} \right. \left\{ \begin{array}{l} Q_1 = -2 + \frac{256 - Y}{16}; \\ Q_2 = -20 - \frac{256 - Y}{16}; \\ Q_3 = 6; \\ Q_4 = -8; \end{array} \right. \left\{ \begin{array}{l} Q_1 = 6; \\ Q_2 = 12; \\ Q_3 = 2 + \frac{Y}{32}; \\ Q_4 = -16 + \frac{Y}{16}. \end{array} \right. \quad (3)$$

В большинстве случаев после бинаризации на изображении образуются помехи в виде бинарного шума. Уменьшить их влияние можно путем применения к бинарному изображению морфологических операций «расширение» и «сужение».

Данный метод сегментации находит области, имеющие цвет, схожий с цветом кожи. Его хорошо применять для локализации рук, лиц и т. д. Однако есть вероятность появления нежелательных объектов (стены, одежда и т. п.), имеющих подобный цвет. Результат работы данного алгоритма представлен на рис. 1.

На основе бинарного изображения, полученного на этапе сегментации, можно построить скелет руки. Для определения скелета используется понятие максимального пустого круга.

Определение 1. Для многоугольной фигуры F максимальным пустым кругом будем называть всякий круг B , полностью содержащийся внутри фигуры F , такой, что любой другой круг B' , содержащийся внутри фигуры F , не содержит в себе B [1].

Используя понятие максимального пустого круга, определим скелет следующим образом.

Определение 2. Скелетом многоугольной фигуры F является множество центров ее максимальных пустых кругов.

На скелете определена радиальная функция $R(x, y)$, ставящая в соответствие каждой точке скелета (x, y) значение радиуса максимального пустого круга с центром в этой точке.

Геометрически скелет многоугольной фигуры представляет собой граф, состоящий из вершин – точек на плоскости и ребер – линий, соединяющих некоторые пары вершин. Причем ребром этого графа может являться либо отрезок, либо дуга параболы [1]. Степень любой вершины в таком графе будет равна 1, 2 или 3.

Существующие эффективные алгоритмы [1; 2] позволяют выполнять построение скелета за время $O(N \log N)$, где N – число вершин в многоугольнике. Построение скелета плоской фигуры подробно описано в статье [2]. В связи с тем, что скорость построения скелета напрямую зависит от количества углов многоугольной фигуры, для ускорения построения скелета можно применить аппроксимацию этой фигуры.

Суть алгоритмов аппроксимации заключается в том, чтобы по заданной многоугольной фигуре получить другую фигуру с меньшим числом вершин. Наиболее популярным и быстрым в исполнении является алгоритм Дугласа–Пекера. Данный алгоритм с указанной точностью позволяет аппроксимировать ломаную или многоугольную фигуру. Алгоритм можно представить в виде нескольких шагов:

Шаг 1. Вычисляется максимальное расстояние между вершинами исходной фигуры. Эти вершины помечаем как M и N .

Шаг 2. Исходная фигура разбивается на две ломаные так, что первая и последняя вершина для каждой ломаной будет лежать на M и N .

Шаг 3. Для каждой ломаной вычисляется максимальное расстояние от отрезка MN до вершины ломаной. Данное расстояние помечаем H , а вершину K .

Шаг 4. Если H меньше или равно заданной точности E , то из текущей ломаной удаляем все вершины, лежащие между M и N . Иначе для ломаных $M...K$ и $K...N$ повторяются аналогичные действия с шага 2 по шаг 4.

На рис. 2 представлена демонстрация алгоритма Дугласа–Пекера и скелеты исходного и аппроксимированного изображения. Внешне исходное и аппроксимированное изображения практически идентичны. Однако контур исходного изображения имеет 74 вершины, а аппроксимированного – 25 вершин, при этом точность аппроксимации 2 пикселя.

После построения скелета обычно выполняется его дополнительная обработка, называемая стрижкой [1], с целью удаления малозначимых и шумовых ветвей.

Демонстрация процесса построения скелета представлена на рис. 3.

Скелет ладони позволяет анализировать направление жеста в пространстве и определить состояние пальцев. Ключевыми точками скелета ладони будем

называть точки начала и конца пальца и точку центра кисти руки.

Каждый палец может принимать два условных состояния: сжатый в кулак или разжатый. Все ветви скелета, соответствующие пальцу, оканчиваются вершиной степени 1. Ветвь пальца можно разделить на две части: палец и пясть. Для классификации ветвей пальцев используется набор эвристических правил:

1. Ветвь пальца лежит на графе между вершинами со степенями 1 и 3.

2. Радиальная функция ветви на вершине степени 1 увеличивается более чем в 2,5 по сравнению с вершиной степени 3.

3. Радиальная функция начинает резко расти, т. е. частные производные R' больше заданного порога (экспериментально было получено число 0,5).

Первая точка на ветви, где производная радиальной функции превышает заданный порог, является точкой конца пальца.

Центром ладони будем считать точку, лежащую на скелете ладони, радиальная функция которой принимает максимальное значение. На рис. 4 демонстрируется результат вычисления ключевых точек на изображении.

Для распознавания простого, ограниченного набора жестов достаточно составить набор эвристических правил, основанных на следующих данных: количество пальцев, их длина, количество циклов в графе и их габариты. В более сложных случаях набора эвристических правил мало и для распознавания жестов применяются дескрипторы формы кисти руки, состоящие из определенных инвариантных признаков.

Дескриптором кисти руки является вектор признаков, позволяющих с определенной точностью определить состояние руки. Дескриптор кисти руки должен быть инвариантен к масштабированию и вращению. Составляющие признаки дескриптора вычисляются из силуэта кисти руки, полученного на этапе сегментации и из его скелета.

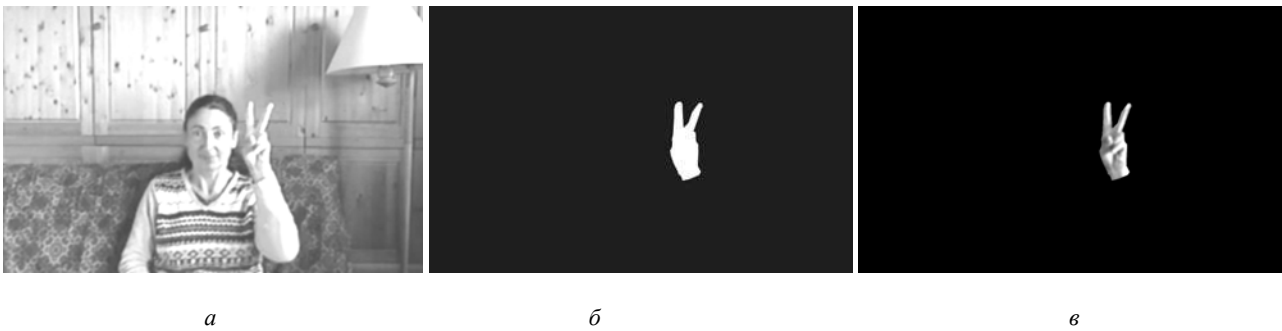


Рис. 1. Результат сегментации изображения: *a* – исходное изображение; *б* – отсегментированное бинарное изображение; *в* – отсегментированное изображение

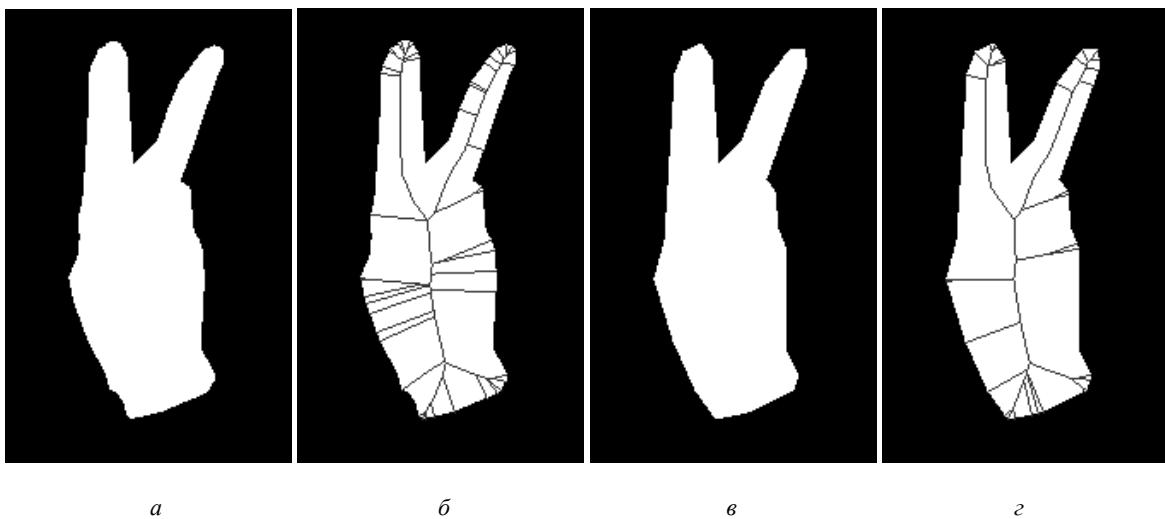


Рис. 2. Скелеты исходного и аппроксимированного изображения: *a* – исходное изображение; *б* – скелет исходного изображения; *в* – аппроксимированное изображение; *г* – скелет аппроксимированного изображения

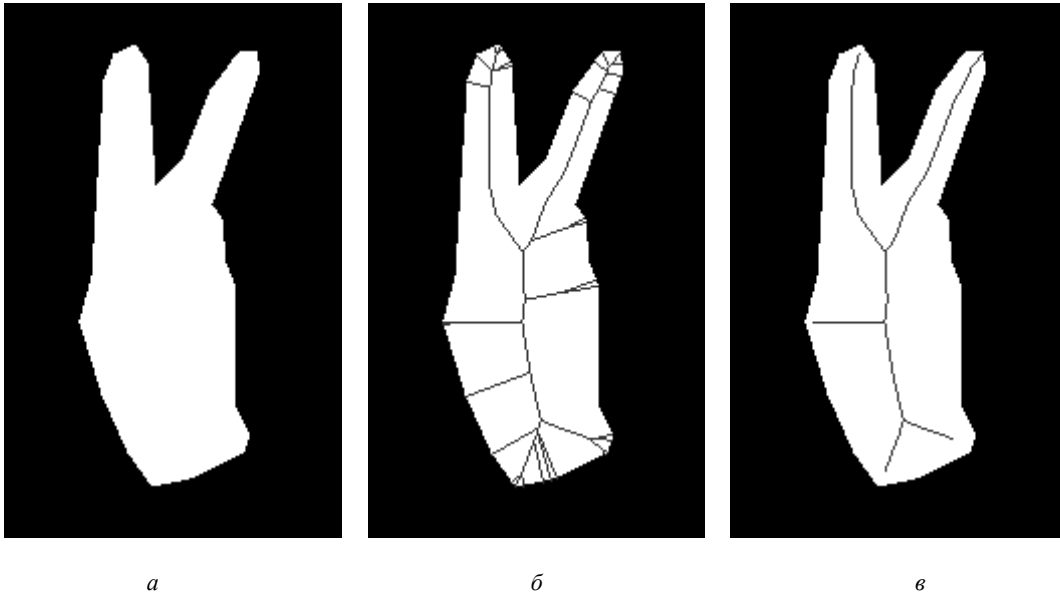


Рис. 3. Процесс построения скелета: *а* – исходное изображение; *б* – скелет исходного изображения; *в* – скелет после стрижки

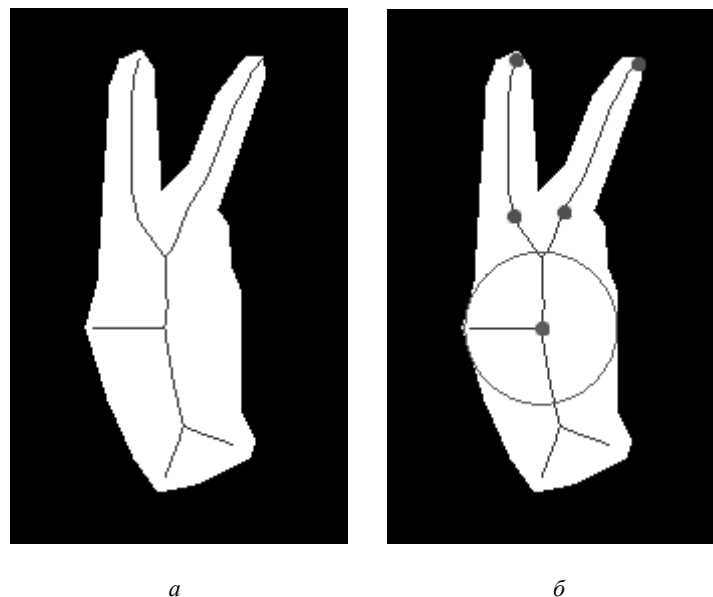


Рис. 4. Определение ключевых точек: *а* – скелет исходного изображения; *б* – скелет и ключевые точки на нем

На основе скелета и его ключевых точек определяется набор параметров:

1. Длина пальцев (длина ветки от начала до конца пальца).
2. Средняя толщина пальцев (среднее значение радиальной функции на ветке скелета, принадлежащей пальцу).
3. Направление пальцев (вектор, начало которого является концом пальца, а конец – началом пальца).
4. Длина циклов скелета.

Распознавание жестов представляет собой поиск наиболее схожих жестов в базе данных. Тип жеста в базе данных, параметры которого наиболее близки

с распознаваемым жестом, будет являться результатом распознавания.

В некоторых случаях два разных жеста могут иметь схожий скелет, поэтому для более точного распознавания предлагается использовать моменты H_u [9]. Моменты H_u инвариантны к масштабированию и вращению и являются хорошим дескриптором для распознавания жестов. Для вычисления моментов H_u необходимо вычислить дискретные, центральные и нормализованные моменты. Дискретные моменты вычисляются на основании силуэта руки, полученного на этапе сегментации. На данном этапе изображение представляет собой двумерную функцию интенсивно-

сти $I(x, y)$, где ненулевые значения пикселей принадлежат кисти руки. Дискретные моменты порядка $(p + q)$ вычисляются по формуле (4), где M и N являются шириной и высотой изображения соответственно:

$$m_{pq} = \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} x^p y^q I(x, y). \quad (4)$$

Центральные моменты инвариантны к перемещению. Моменты порядка $(p + q)$ вычисляются на основе дискретных моментов по формуле (5), где \bar{x} и \bar{y} являются центром тяжести отсегментированного изображения:

$$\begin{aligned} \mu_{pq} &= \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} (x - \bar{x})^p (y - \bar{y})^q I(x, y), \\ \bar{x} &= m_{10} / m_{00} \quad \bar{y} = m_{01} / m_{00}. \end{aligned} \quad (5)$$

Точка с координатой (\bar{x}, \bar{y}) является центром масс изображения. Нормализованные моменты являются инвариантными к масштабированию и вычисляются по формуле (6), где $\gamma = 1 + (p + q) / 2$:

$$\eta_{pq} = \mu_{pq} / \mu_{pq}^\gamma. \quad (6)$$

Семь значений моментов H_u ϕ_1 – ϕ_7 вычисляются по формулам (7). Моменты H_u инвариантны к вращению и масштабированию:

$$\begin{aligned} \phi_1 &= \eta_{20} - \eta_{02}; \\ \phi_2 &= (\eta_{20} - \eta_{02})^2 + 4\eta_{11}^2; \\ \phi_3 &= (\eta_{30} - 3\eta_{12})^2 + (3\eta_{21} - \eta_{03})^2; \\ \phi_4 &= (\eta_{30} + \eta_{12})^2 + (\eta_{21} + \eta_{03})^2; \\ \phi_5 &= (\eta_{30} - 3\eta_{12})(\eta_{30} + \eta_{12}) \left[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2 \right] + \\ &\quad + (3\eta_{21} - \eta_{03})(\eta_{21} + \eta_{03}) \left[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2 \right]; \\ \phi_6 &= (\eta_{20} - \eta_{02}) \left[(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2 \right] + \\ &\quad + 4\eta_{11}(\eta_{30} + \eta_{12})(\eta_{21} + \eta_{03}); \\ \phi_7 &= (3\eta_{21} - \eta_{03})(\eta_{30} + \eta_{12}) \left[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2 \right] - \\ &\quad - (\eta_{30} - 3\eta_{12})(\eta_{21} + \eta_{03}) \left[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2 \right]. \end{aligned} \quad (7)$$

Для проведения исследований был разработан программный продукт, позволяющий локализовать кисть руки методом цветовой сегментации и распознавать жесты на основе скелетной модели руки. В таблице приведены результаты истинного и ложного распознавания жестов рук на основе модели непрерывного скелета, на основе аппроксимированного скелета и на основе аппроксимированного скелета + моментов H_u . Изображения тестировались на компьютере с процессором Intel(R) Core(TM) i5 750 2.67 GHz. При тестировании использовалась база динамических жестов соревнования ChaLearn Gesture Challenge [3]. Данная база состоит из набора независимых пакетов. Каждый из пакетов содержит по 10 обучающих жестов и 30–40 контрольных видео по несколько жестов в каждом. Все изображения были размером 230×240 пикселей.

Оценка методов распознавания жестов рук

№	Непрерывный скелет			Аппроксимированный скелет			Аппроксимированный скелет + моменты H_u		
	Истина, %	Ложь, %	Время, мс	Истина, %	Ложь, %	Время, мс	Истина, %	Ложь, %	Время, мс
“0”	53,5	32,5	6,8	59,4	26,6	7,2	87,1	5	10,5
“1”	60,8	28,1	7	66,6	21,5	7,4	92,9	3,3	9,2
“2”	53,7	28,9	6,3	59,2	23,9	6,6	86,7	3,1	8,6
“3”	57,7	33,8	6,9	64,5	28,3	7,4	93,2	7,7	10,4
“4”	58,8	28,6	4,1	64,3	21,8	4,6	90,4	2,3	7,7
“5”	58,7	31,5	4,7	65,4	24,9	5	93,5	3,9	8
“6”	60,4	30,3	6	66,4	23,9	6,5	93,3	6,6	9,3
“7”	53,6	31,8	5	58,9	25,6	5,6	87,9	4,3	7,2
“8”	56,4	33,8	5,9	63,1	28,6	6,4	87,9	8	8,7
“9”	53	30	3,7	59,2	25	4,1	87,8	6,8	6,2

Библиографические ссылки

1. Местецкий Л. М. Непрерывная морфология бинарных изображений: фигуры, скелеты, циркуляры. М.: Физматлит, 2009.
 2. Местецкий Л. М., Рейер И. Непрерывное скелетное представление изображения с контролируемой

точностью // International Conference Graphicon. М., 2003. С. 51–54.
 3. ChaLearn Gesture Dataset (CGD2011). ChaLearn, California, 2011 [Электронный ресурс]. URL: <http://gesture.chalearn.org>.
 4. Gudmundsson S. A. [et al.] Model-Based Hand Gesture Tracking in ToF Image Sequences // 6th Intern.

Conf. on Articulated motion and deformable objects (AMDO), 2010. P. 118–127.

5. Suryanarayan P., Subramanian A., Mandalapu D. Dynamic Hand Pose Recognition Using Depth Data // 20th Intern. Conf. on Pattern Recognition (ICPR), 2010. P. 3105–3108.

6. Phung S. L., Bouzerdoum A., Chai D. Skin Segmentation Using Color Pixel Classification: Analysis and Comparison // IEEE Trans. Pattern Anal. Mach. Intell. 2005. Vol. 27. P. 148–154.

7. Siddharth J., Gaurav S., Face detection // EE368: Digital Image Processing. 2003. P. 101–112.

8. Зотин А. Г., Носов А. В., Бузаев Д. В. Анализ пригодности методов сегментации для локализации объектов на основе цветовых и структурных признаков // Вестник СибГАУ. 2012. Вып. 1(41). С. 23–28.

9. Hu MK. Visual Pattern Recognition by Moment Invariant. IRE Trans. Info. 1962. Theory 8 (2). P. 179–187.

References

1. Mesteckij L. M. *Nepřerývnaja morfologija binarnyh izobrazhenij: figury, skelety, cirkuljary* [Continuous morphology of binary images : figures, skeletons, circulars]. Moscow, Fizmatlit Publ., 2009.

2. Mesteckij L. M., Rejer I. [Continuous skeletal representation of the image with controlled accuracy]. *International Conference Graphicon*. Moscow, 2003, p. 51–54.

3. ChaLearn Gesture Dataset (CGD2011). ChaLearn, California, 2011, Available at URL: <http://gesture.chalearn.org> (accessed 5 February 2014).

4. Gudmundsson S. A., Sveinsson J. R., Pard'as M. et al. Model-Based Hand Gesture Tracking in ToF Image Sequences. 6th International Conference on Articulated motion and deformable objects (AMDO), 2010, p. 118–127.

5. Suryanarayan P., Subramanian A., Mandalapu D. Dynamic Hand Pose Recognition Using Depth Data. 20th International Conf. on Pattern Recognition (ICPR), 2010, p. 3105–3108.

6. Phung S. L., Bouzerdoum A., Chai D. Skin Segmentation Using Color Pixel Classification: Analysis and Comparison. *IEEE Trans. Pattern Anal. Mach. Intell.* 2005. January. Vol. 27, p. 148–154.

7. Siddharth J., Gaurav S. Face detection. *EE368: Digital Image Processing*. 2003, p. 101–112.

8. Zotin A. G., Nosov A. V., Buzaev D. V. [Suitability analysis segmentation methods for localization of objects based on the color and structural features]. *Vestnik SibGAU*. 2012, vol. 41, no. 1, p. 23–28. (In Russ.)

9. Hu MK. Visual Pattern Recognition by Moment Invariant. IRE Trans. Info. 1962, Theory 8 (2), p. 179–187.

© Носов А. В., 2014

УДК 539.3

К РАСЧЕТУ ПЛАСТИН В УСЛОВИЯХ ПЛОСКОГО НАПРЯЖЕННОГО СОСТОЯНИЯ НА ТЕМПЕРАТУРНЫЕ НАГРУЗКИ ВАРИАЦИОННО-РАЗНОСТНЫМ МЕТОДОМ В ФУНКЦИЯХ НАПРЯЖЕНИЙ

Р. А. Сабиров

Сибирский государственный аэрокосмический университет имени академика М. Ф. Решетнева
Российская Федерация, 660014, г. Красноярск, просп. им. газ. «Красноярский рабочий», 31
E-mail: rashidsab@mail.ru

Разработан вариационно-разностный метод расчета напряженного состояния пластинок в условиях плоской задачи теории упругости в функциях напряжений. Для решения температурной задачи применяется метод устранения деформаций С. Тимошенко. Получен функционал в функциях напряжений с учетом температурного члена, что освобождает от вычисления перемещений и деформаций. Особенностью разработанного алгоритма расчета является использование для формирования коэффициентов системы разрешающих уравнений и ее правой части первой и второй вариаций данного функционала. Это дает простой и универсальный алгоритм вычислений – применяются одни и те же процедуры в программных модулях. Конечно-разностная дискретизация континуальной задачи позволяет решать задачи большой размерности. Напряжения на контуре известны априори; в области пластинки уравнения неразрывности обеспечиваются при достаточно редких конечно-разностных сетках, экономится время счета и ресурсы оперативной памяти. Составлена программа расчета на основе пакета Maple; приведен пример расчета свободной пластинки при неравномерном нагреве.

Ключевые слова: плоская задача теории упругости, функционал Кастилиано, вариационно-разностный метод, функция напряжений.