

УДК 681.3.15

МЕТОД РАСПОЗНАВАНИЯ ЛИЦ С ПРИМЕНЕНИЕМ СКАЛЯРНЫХ ФУНКЦИЙ ВОЗМУЩЕНИЯ И ТЕОРЕТИКО-МНОЖЕСТВЕННОЙ ОПЕРАЦИИ ВЫЧИТАНИЯ

С. И. Вяткин

*Институт автоматизи и электрометрии СО РАН,
630090, г. Новосибирск, просп. Академика Коптюга, 1
E-mail: sivser@mail.ru*

Обсуждается проблема идентификации лиц. Предложен метод их распознавания на основе скалярных функций возмущения и теоретико-множественной операции вычитания. Показано, что в процессе тестирования образцов используются все точки поверхностей и объём масок для более точной идентификации.

Ключевые слова: стереопара, карта глубины, алгоритм корреляции, функции возмущения, операция вычитания, вокселизация.

DOI: 10.15372/AUT20160107

Введение. Исследования по идентификации лиц ведутся уже давно. Ранее такая задача основывалась на двухмерных ($2D$) изображениях: их получении и сравнении с имеющейся базой данных [1]. Однако в большинстве случаев незначительные изменения положения объекта наблюдения делали работу системы распознавания неэффективной. Тестируемые образцы не удавалось сопоставить ни с одной из баз данных, и частота ошибок оказывалась слишком высокой.

Известны следующие методы идентификации лиц. Лицо человека снимает видеокамера, и изображение обрабатывается специальными фильтрами. Далее автоматически решается задача выделения на изображении особых точек FFE (Face Feature Extraction), после чего из этих точек (и расстояний между ними) образуется эталон, с которым и производится сравнение. К достоинствам метода относится возможность осуществления непрерывной идентификации скрытно. Недостатки метода — зависимость его от поворота головы и внешних особенностей лица.

Альтернативой видеоизображениям является термограмма, которая регистрируется инфракрасной камерой, и изображение обрабатывается. Метод удобен тем, что человека можно регистрировать в полной темноте, и это повышает скрытность процесса. Однако появляется зависимость от внешних источников тепловых помех, кроме того, требуется использование специального оборудования.

Трёхмерное ($3D$) распознавание — один из наиболее прогрессивных методов [2], суть которого состоит в следующем. На лицо проецируются линии, по которым восстанавливается его $3D$ -модель. В этой модели выделяются специальные точки, образующие вектор признаков. Достоинства метода — непрерывная и скрытная идентификация объекта; невозможно использовать муляж; различаются близнецы; малая зависимость от поворота головы (диапазон отклонения значительно увеличен); кроме того, при правильном выборе светового диапазона малая зависимость от внешней освещённости, от волосяного покрова и опухлостей лица. Трёхмерная идентификация может быть использована в тёмной среде, остаётся эффективной даже при повороте головы до 90° . Недостатки метода — необходимость специального оборудования, высокие вычислительно-мощностные требования (аппаратная реализация алгоритмов), что увеличивает стоимость системы. Система

распознавания в процессе идентификации производит ряд действий. Изображение лица может быть получено путём цифрового сканирования существующей 2D-картинки или использования видеоизображения. Когда лицо обнаружено, происходит его «выравнивание», т. е. система определяет положение и размер головы. Как упоминалось ранее, лицо распознаётся при повороте головы до 90° , тогда как при 2D-идентификации допустимый угол поворота не более 35° . При измерении лица система вычисляет кривые линии менее чем в миллиметровом масштабе и создаёт шаблон, а программа переводит шаблон в числовой код. Таким образом, каждое лицо в итоге представляется в цифровом виде. Далее происходит процесс сопоставления изображений. Все лица обладают рядом отличительных черт, выступов и ямочек, которые делают их уникальными. Ключевые черты фиксируются в качестве узлов — таких у человека в среднем порядка 80. Расстояния между ними используются программой для сравнения разных лиц. Наиболее важные из расстояний: промежуток между глазами, ширина носа, глубина глазниц, форма скуловых костей, длина линии челюсти.

В данной работе представлен метод распознавания лиц с помощью скалярных функций [3] возмущения и теоретико-множественной операции вычитания [4].

Описание метода. Для вычисления трёхмерных точек лица используется калиброванная стереопара (рис. 1). Пусть имеются две проективные матрицы M_i :

$$\begin{pmatrix} u_i s_i \\ v_i s_i \\ s_i \end{pmatrix} = M_i \begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix}, \quad (1)$$

где x, y, z — трёхмерные координаты точки; u_i и v_i — их проекции на изображениях i ; s_i — масштабный множитель. Стереопару составляют следующие параметры: $E_1 = (u_1, v_1)$, $E_2 = (u_2, v_2)$ — точки плоскостей изображения; $P = (x, y, z)$ — точка мировой системы координат.

Из калиброванной стереопары лица вычисляется карта глубины (рис. 2) с применением алгоритма корреляции. Существуют два основных вида алгоритмов вычисления трёхмерных данных по стереоизображению: на основе особенностей (Feature-based) и площади (Area-based) изображений [5–7].

Алгоритмы на основе особенностей выделяют интересующие элементы (рёбра, сегменты, контуры) на изображениях и подбирают их для двух и более точек наблюдения.

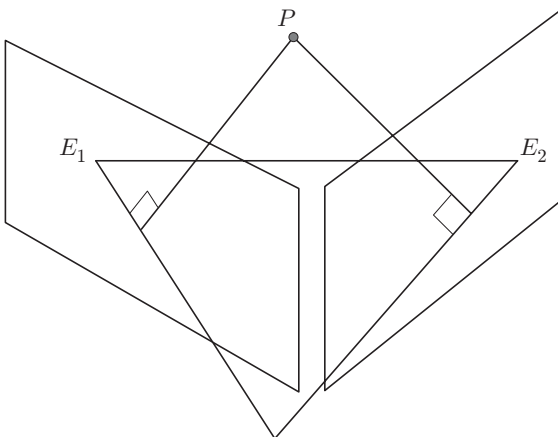


Рис. 1

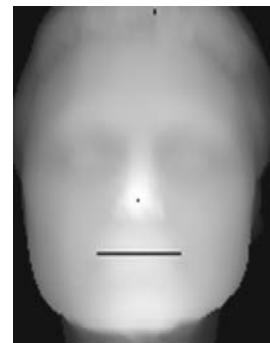


Рис. 2

Достоинства: алгоритмы быстрые, используется небольшое число пикселей. Недостатки: не гарантируют удачного поиска выбранных примитивов.

Поэтому в данной работе применяется алгоритм на основе площади с корреляцией уровней интенсивности изображения [5]

$$s = \frac{\sum_{i,j} ((I_1(x+i, y+j) - \bar{I}_1) - (I_2(x+dx+i, y+dy+j) - \bar{I}_2))^2}{\sqrt{\left(\sum_{i,j} (I_1(x+i, y+j) - \bar{I}_1)^2\right) \left(\sum_{i,j} (I_2(x+dx+i, y+dy+j) - \bar{I}_2)^2\right)}}, \quad (2)$$

где I_1, I_2 — интенсивности левого и правого изображений; \bar{I}_1, \bar{I}_2 — их средние значения; dx, dy — смещения вдоль эпиполярной линии; $s = \max(0, 1 - c)$ — оценка корреляции.

Имеются два изображения стереопары (см. рис. 1), при их сканировании определяются данные буфера глубины (карта высот) в соответствии с (2) и алгоритмом, описанными в работе [5].

Карты уровней детальности вычисляются с использованием данных буфера глубины [3]. Открытое односвязное множество точек на плоскости будем называть плоской областью. Пусть D — плоская область, а \bar{D} — её замыкание. Введём на плоскости координатную систему (u, v) . Пусть x, y, z — прямоугольные декартовы координаты точек в трёхмерном евклидовом пространстве E^3 . Зададим на множестве \bar{D} три непрерывные функции:

$$x = \varphi(u, v); \quad y = \psi(u, v); \quad z = \chi(u, v). \quad (3)$$

Предположим, что функции (3) обладают следующими свойствами. Если (u_1, v_1) и (u_2, v_2) — различные точки множества \bar{D} , то точки $M_1(x_1, y_1, z_1)$ и $M_2(x_2, y_2, z_2)$ пространства E^3 , координаты которых вычислены по формулам (3), тоже различны:

$$\begin{aligned} x_1 = \varphi(u_1, v_1); \quad y_1 = \psi(u_1, v_1); \quad z_1 = \chi(u_1, v_1); \\ x_2 = \varphi(u_2, v_2); \quad y_2 = \psi(u_2, v_2); \quad z_2 = \chi(u_2, v_2). \end{aligned} \quad (4)$$

Множество S точек $M(x, y, z)$, координаты x, y, z которых определяются соотношениями $x = \varphi(u, v)$, $y = \psi(u, v)$, $z = \chi(u, v)$, где функции φ, ψ, χ в замыкании \bar{D} области D обладают вышеописанным свойством, является простой поверхностью.

Далее строится сложная поверхность (слепок) F , т. е. график функции, определённой в $3D$ -пространстве $z = f(x, y)$ (рис. 3). Задание сложной поверхности на основе скалярного поля представляет собой совокупность некой базовой поверхности P , находящейся в той же системе координат, что и F , и связанной с базовой поверхностью P карты высот.

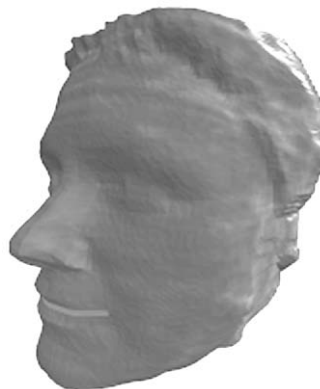


Рис. 3

Карта высот представляет собой $2D$ -прямоугольник, называемый областью возмущения D_P базовой поверхности P , внутри которого задана функция возмущения $h(u, v)$. Карта высот, в свою очередь, является самым возмущением [3]. Область определения функции $h(u, v)$ есть $D_{h(u, v)} = \{U, V\}$, где U и V — размеры прямоугольника. Карта высот связана с базовой поверхностью следующим образом: существует преобразование $G(\mathbb{R}^3 \Rightarrow \mathbb{R}^2)$ из системы координат, в которой находятся F и P , в систему координат карты высот. Как правило, таким преобразованием будет параллельная проекция.

Значение функции $h(G(\mathbf{d}_F))$ характеризует отклонение точки \mathbf{d}_F , лежащей на поверхности F , от точки \mathbf{d}_P , являющейся проекцией этой точки на поверхность P . Другими словами, значение функции $h(G(\mathbf{d}_F))$ равно модулю вектора

$$\mathbf{v} = (\mathbf{d}_F - \mathbf{d}_P). \quad (5)$$

Следовательно, область сложной поверхности можно принять как множество точек в \mathbb{R}^3 , задаваемых векторным уравнением

$$\mathbf{F} = G(\mathbf{v}) + \mathbf{n}h(G(\mathbf{v})); \quad \forall \mathbf{v} \in \mathbb{R}^3, \quad (6)$$

где \mathbf{n} — нормаль к базовой поверхности.

Если \mathbf{v} находится за пределами области возмущения, то вектор $\mathbf{n}h(G(\mathbf{v})) = 0$, а вектор \mathbf{F} является вектором на базовой поверхности. Таким образом, для задания формы возмущающей поверхности можно использовать таблицу чисел из [3], а в качестве функции h — функцию интерполяции по узловым значениям, которые берутся из этой же таблицы.

При определении максимума возмущения вычисляется характерный размер проекции текущего интервала, в зависимости от которого выбирается уровень детальности. Для большего интервала соответственно берётся грубое приближение исходной функции. Если требуется более точное представление, чем имеется, то производится билинейная или бикубическая интерполяция значений высот, составляющих последний уровень детальности. В результате получается объёмный трёхмерный слепок лицевой части (см. рис. 3). По трём антропоморфным точкам строится система координат для возможности совмещения тестируемых слепков (рис. 4), по завершении которого клипширующей плоскостью отсекаются части для выравнивания объёмов.

После теоретико-множественной операции вычитания (рис. 5)

$$\mathbf{F}_3 = \mathbf{F}_1 \setminus \mathbf{F}_2 \quad (7)$$

определяется количество $3D$ -точек (вокселей), принадлежащих объекту $f_3 = \Phi_i(f_1(x, y, z), f_2(x, y, z))$, $\mathbf{F}_1: f_1(x, y, z) \geq 0$, $\mathbf{F}_2: f_2(x, y, z) \geq 0$. Чтобы найти трёхмерные точки, необходимо вокселизировать оставшуюся часть объёма после вычитания [8]. Исходя из этого, будем

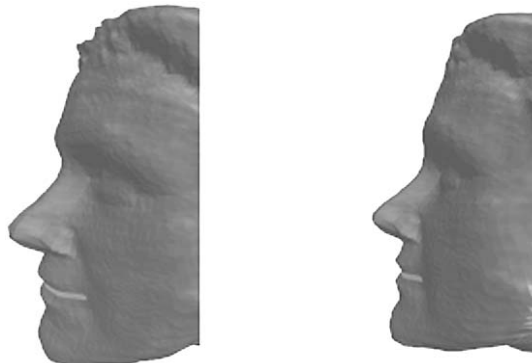


Рис. 4

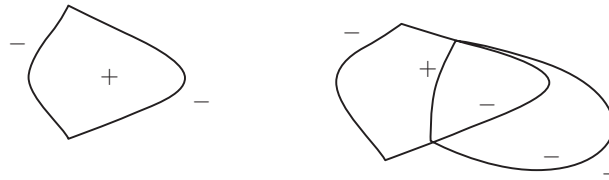


Рис. 5

считать, что наблюдатель смотрит вдоль оси Z . Требуется получить проекцию оставшейся части на плоскость XY , которая должна представлять собой конечный набор значений. Поэтому весь куб будет делиться на «бруски» так, чтобы каждый брусок соответствовал пикселю на изображении. Каждый брусок будет делиться вдоль оси Z , образуя набор вокселей. Осуществляется бинарное деление бруска по z -координате (назовём его вокселем верхнего уровня). Воксель на уровне i разбивается на две равные части (на два вокселя следующего уровня детальности). Далее каждая часть рекурсивно разбивается в поиске вокселей, принадлежащих рассматриваемому остатку. Вектор ближнего по отношению к наблюдателю конца вокселя равен полусумме векторов ближнего и дальнего концов вокселя предыдущего уровня деления, а вектор дальнего конца вокселя равен вектору дальнего конца вокселя предыдущего уровня деления:

$$\begin{aligned}
 P_{ni}^1 &= P_{ni-1}; & P_{fi}^1 &= (P_{ni-1} + P_{fi-1})/2; \\
 P_{ni}^2 &= (P_{ni-1} + P_{fi-1})/2; & P_{fi}^2 &= P_{fi-1}; \\
 V_i^1 &= \{P_{ni}^1, P_{fi}^1\}; & V_i^2 &= \{P_{ni}^2, P_{fi}^2\},
 \end{aligned}
 \tag{8}$$

где V_i^1, V_i^2 — воксели i -го уровня рекурсии (ближний к точке и дальний от точки наблюдения соответственно); P_{ni}^1 и P_{fi}^2 — координаты ближнего по отношению к наблюдателю и дальнего концов вокселя на i -м уровне рекурсии.

Чем меньше вокселей остаётся, тем больше сходство тестируемых объектов.

Анализ метода и результаты работы. При тестировании метода результаты оказались положительными. Применялись как виртуальные объекты из доступных баз данных [9–11], так и реальные персонажи. Для конвертирования полигональных моделей в формат скалярных функций возмущения использовался метод, представленный в [12].

Трёхмерная технология распознавания лиц эффективно работает, с её помощью были успешно распознаны более 98 % тестовых объектов. Однако существует ряд причин, приводящих к неудачной верификации. Их можно разделить на две категории: неправильное расположение перед камерой и помехи при считывании данных. К первой категории относятся ситуации, когда в область обзора камеры попадает только часть лица: лицо человека направлено не на камеру, голова наклонена вниз или в сторону от камеры, человек стоит слишком близко к камере, а также отходит от камеры слишком быстро после начала верификации (менее секунды). Метод успешно работает, когда распознаваемый объект передвигается равномерно, но если он ускоряется резко, камера не успевает точно фиксировать объект наблюдения.

Следует заметить, что частичное попадание лица в обзор камеры не является полностью неприемлемым, так как фрагменты могут успешно проверяться с помощью геометрической операции пересечения [4]. В предлагаемом методе возможно избирательное тестирование с применением геометрической операции пересечения прозрачного цилиндра или любой другой формы с поверхностью.

Помехи при считывании данных возникают при несоблюдении требования сохранения нейтрального выражения лица или когда, например, головной убор, солнцезащитные зеркальные очки или волосы закрывают большую часть лица.

Современные методы могут распознавать персонажи по разным выражениям лица [2, 13].

В предлагаемом методе для распознавания используется трёхмерный морфинг [4].

Если сравнивать двухмерные системы и предлагаемый трёхмерный метод распознавания, то в первом случае при установке порога распознавания в 70 % уровень вероятности ложного срабатывания составляет 0,12 %, а уровень вероятности ложного отказа — 9,8 % [1, 14]. Во втором случае порог распознавания был установлен в 90 %, удалось достигнуть уровня вероятности ложного срабатывания 0,004 % и уровня вероятности ложного отказа 0,1 %.

По всем тестам, проводившимся одновременно для обеих технологий с помощью одних и тех же изображений, трёхмерная технология распознавания лиц оказалась эффективнее двухмерной.

Примером 3D-методов распознавания является хорошо известный метод выравнивания для восстановления формы и параметров текстуры [15, 16]. Этот метод использует систему линейных уравнений. Распознавание проводится путём сопоставления восстановленных форм и текстуры изображения.

Однако в методе [15, 16] необходима ручная инициализация. Время распознавания около 1 мин на процессоре Pentium III с частотой 800 МГц не удовлетворяет требованиям большинства реальных систем.

По сравнению с известными методами предлагаемый имеет следующие преимущества: трёхмерный морфинг решает задачу распознавания лица по разным выражениям; возможна идентификация лица по части изображения; не надо текстурировать поверхность лица; является полностью автоматическим и имеет быструю скорость (около 200 мс на изображение лица с разрешением 640×480 пикселей с применением процессора Intel Core i7-2700K (8 Мбайт кэш-памяти, 3,90 ГГц)), что примерно на два порядка быстрее, чем в методе выравнивания [15, 16]. Точность измерения составляет не более 0,8 мм (для каждой точки трёхмерной поверхности).

Для визуализации в реальном времени при необходимости используется бинарный метод поиска элементов изображения с помощью графических ускорителей [17], адаптированный для вычисления скалярных функций возмущения [18, 19]. В методе отображения поверхности и смены уровней детальности в основе лежит тот же механизм, что и для обычной текстуры. Для фильтрации изображения в динамике применяется интерполяция высот по аналогии с текстурой цвета. Время вычислений практически не зависит от разрешения карты высот, а зависит от разрешения экрана (квадратичная зависимость) и от разрешения по глубине Z (логарифмическая зависимость). Кроме того, не требуется триангуляция поверхности. Эта технология основана на иерархическом механизме управления уровнями детальности, который напоминает MIP-mapping обычной текстуры цвета. Если проводить полную аналогию с текстурой, то все её положительные свойства наследуются данной технологией, в том числе и простота анимации такой поверхности без вычислительных затрат на геометрические преобразования вершин треугольников в отличие от поверхности, аппроксимированной треугольниками.

Заключение. В представленной работе предложен метод распознавания лиц на основе скалярных функций возмущения и теоретико-множественной операции вычитания. Для распознавания лиц использовались трёхмерные слепки. Данный метод отличается от известных трёхмерных методов тем, что в процессе проверки участвуют не только все точки поверхности, но и объём тестируемой маски. Метод имеет следующие преимущества: не

требуется ручной инициализации процесса; трёхмерный морфинг решает задачу распознавания лица по разным выражениям; возможна идентификация лица по части изображения; метод реконструкции лица является полностью автоматическим. Время вычислений составляет около 200 мс при разрешении 640×480 пикселей.

Метод возможно применять в областях, где требуется интеллектуальный видеоконтроль особо охраняемых объектов: предприятий оборонного комплекса, мест большого скопления людей и других.

СПИСОК ЛИТЕРАТУРЫ

1. **Abate A. F., Nappi M., Riccio D., Sabatino G.** 2D and 3D face recognition: a survey // *Pattern Recogn. Lett.* 2007. **28**, N 14. P. 1885–1906. DOI: 10.1016/j.patrec.2006.12.018.
2. **Kakadiaris I. A., Passalis G., Toderici G. et al.** Three-dimensional face recognition in the presence of facial expressions: an annotated deformable model approach // *IEEE Trans. Pattern Anal. Machine Intell.* 2007. **29**, N 4. P. 640–649. DOI: 10.1109/TPAMI.2007.1017.
3. **Вяткин С. И.** Моделирование сложных поверхностей с применением функций возмущения // *Автометрия.* 2007. **43**, № 3. С. 40–47.
4. **Вяткин С. И.** Преобразования функционально заданных форм // *Программные системы и вычислительные методы.* 2014. № 4. С. 484–499.
5. **Fua P.** A parallel stereo algorithm that produces dense depth maps and preserves image features // *Mach. Vis. Appl.* 1993. **6**, N 1. P. 35–49.
6. **Hannah M. J.** Digital stereo image matching techniques // *Intern. Arch. Photogrammetry Remote Sensing.* 1988. **27**, Pt. B3. P. 280–293.
7. **Anandan P.** A computational framework and an algorithm for the measurement of visual motion // *Intern. Journ. Comput. Vis.* 1989. **2**, N 3. P. 283–310.
8. **Катасонов А. В., Вяткин С. И., Долговесов Б. С.** Вокселизация функциональных форм // *Тр. 15-й Междунар. конф. по компьютерной графике и ее приложениям «Графикон-2005».* Новосибирск: ИВММГ СО РАН, 2005. С. 372–377.
9. **The Face Datasets.** URL: <http://vasc.ri.cmu.edu/idb/html/face/> (дата обращения: 24.11.2014).
10. **The Bosphorus Database.** URL: <http://bosphorus.ee.boun.edu.tr/HowtoObtain.aspx> (дата обращения: 24.11.2014).
11. **Face Recognition Homepage. Databases.** URL: <http://www.face-rec.org/databases/> (дата обращения: 24.11.2014).
12. **Vyatkin S. I., Romanyuk S. A., Dudnyak O. O.** Geometric modeling with scalar perturbation functions // *Intern. Sci.-Techn. Magazine "Measuring and Computing Devices in Technological Processes"*. 2014. N 4(49). P. 45–50. URL: http://journals.khnu.km.ua/vottp/pdf/2014_4/7.pdf (дата обращения: 24.11.2014).
13. **Fang T., Zhao X., Ocegueda O. et al.** 3D/4D facial expression analysis: an advanced annotated face model approach // *Image Vis. Comput.* 2012. **30**, N 10. P. 738–749. DOI: 10.1016/j.imavis.2012.02.004.
14. **Toderici G., Passalis G., Zafeiriou S. et al.** Bidirectional relighting for 3D-aided 2D face recognition // *Proc. IEEE Comput. Soc. Conf. Computer Vision and Pattern Recognition.* San Francisco, USA, 2010. P. 2721–2728. DOI: 10.1109/CVPR.2010.5539995.
15. **Blanz V., Romdhani S., Vetter T.** Face-identification across different poses and illuminations with a 3D morphable model // *Proc. of the 5th IEEE Intern. Conf. on Automatic Face and Gesture Recognition.* Washington, USA, 2002. URL: http://gravis.cs.unibas.ch/publications/sgd_fg02.pdf (дата обращения: 24.11.2014).

16. **Romdhani S., Blanz V., Vetter T.** Face identification by fitting a 3D morphable model using linear shape and texture error functions // Proc. of the 7th European Conference on Computer Vision (ECCV'02). Copenhagen, Denmark, May 28–31, 2002. Pt. 4. P. 3–19. URL: http://mi.informatik.uni-siegen.de/publications/romdhani_eccv02.pdf (дата обращения: 24.11.2014).
17. **Вяткин С. И.** Метод бинарного поиска элементов изображения функционально заданных объектов с применением графических акселераторов // Автометрия. 2014. **50**, № 6. С. 89–96.
18. **Vyatkin S. I.** Three-dimensional modeling of the volcano Bandai using shape texture // Intern. Journ. Natur. Sci. Research. 2015. **3**, N 2. P. 21–29. URL: [http://www.pakinsight.com/pdf-files/78-IJNSR-2015-3\(2\)-21-29.pdf](http://www.pakinsight.com/pdf-files/78-IJNSR-2015-3(2)-21-29.pdf) (дата обращения: 24.11.2014).
19. **Вяткин С. И.** Визуализация фотореалистичного рельефа местности на основе текстуры-формы с использованием графических ускорителей // Программные системы и вычислительные методы. 2015. № 1(10). С. 89–107.

Поступила в редакцию 24 ноября 2014 г.
