

ВИЗУАЛИЗАЦИЯ ТРЕХМЕРНЫХ ВИДЕОПОСЛЕДОВАТЕЛЬНОСТЕЙ ИЗОБРАЖЕНИЙ В РЕАЛЬНОМ ВРЕМЕНИ НА ОСНОВЕ МНОГОУРОВНЕВОЙ ДЕКОМПОЗИЦИИ ВЕЙВЛЕТОВ

© 2011 г. В. Ф. Кравченко, В. И. Пономарев, академик В. И. Пустовойт

Поступило 26.05.2011 г.

1. Впервые на основе идей [1–9] предложен и обоснован метод, позволивший реализовать визуализацию трехмерных (3D) видеопоследовательностей изображений на основе двумерных видеопоследовательностей (2D) в реальном времени. Новый подход использует многоуровневую декомпозицию вейвлетов, в частности вейвлетов, основанных на атомарных функциях, позволяя реализовать лучшие аппроксимационные характеристики последних в 3D-визуализации. Процедура, основанная на многоуровневой декомпозиции вейвлетов, базирующаяся на атомарных функциях (WAF) [7, 8], сравнивается с другими известными алгоритмами L&K [3], SSD [4], GEEMSF [5] и классическими вейвлетами (Daubechies, Symlets, Biorthogonal и Coiflets). Разработанный метод основан на восстановлении изображений (карт) глубин с последующей 3D-визуализацией, используя технику формирования анаглифа по двум соседним изображениям (кадрам) 2D-видеопоследовательности. Реализация предложенных алгоритмов 3D-визуализации в реальном масштабе времени была проведена на основе процессора EVM DM642 фирмы “Texas Instruments” для регистрации и обработки изображений.

2. Новый метод включает следующие операции: декомпозицию 2D-видеопоследовательности в кадры изображений; разделение цветного изображения на RGB-компоненты; вычисление глубинных карт (DM) на основе стереопар, формируемых соседними кадрами, с использованием многоуровневой декомпозиции вейвлетов WAF (M-WAF); коррекция глубинной карты с помощью динамического сжатия; синтезирование анаглифа путем интерполяции по методу

ближайшего соседнего пикселя (NNI) и, наконец, визуализация 3D-видеопоследовательности изображений. Блок-схема предлагаемого алгоритма представлена на рис. 1.

Многоуровневая декомпозиция вейвлет-функции формируется на основе дискретного преобразования вейвлета (DWT) для трех компонент деталей [6]: горизонтальной, вертикальной, диагональной, и представляется в виде

$$WT_s = W_s \angle \Theta_{w_s}, \quad W_s = \frac{\sqrt{|D_{h_s}|^2 + |D_{v_s}|^2 + |D_{d_s}|^2}}{|A_s|}, \quad (1)$$

где W_s – нормализованный модуль для вейвлета на декомпозиционном уровне s ; D_{h_s} , D_{v_s} , D_{d_s} – горизонтальная, вертикальная и диагональная компоненты для деталей изображения на том же уровне s ; A_s – это аппроксимационная компонента; Θ_{w_s} – фаза, которая вычисляется так:

$$\Theta_{w_s} = \begin{cases} \alpha_s & \text{в случае } D_{h_s} > 0, \\ \pi - \alpha_s & \text{в случае } D_{h_s} < 0, \end{cases} \quad (2)$$

$$\alpha_s = \arctg \frac{D_{h_s}}{D_{v_s}}.$$

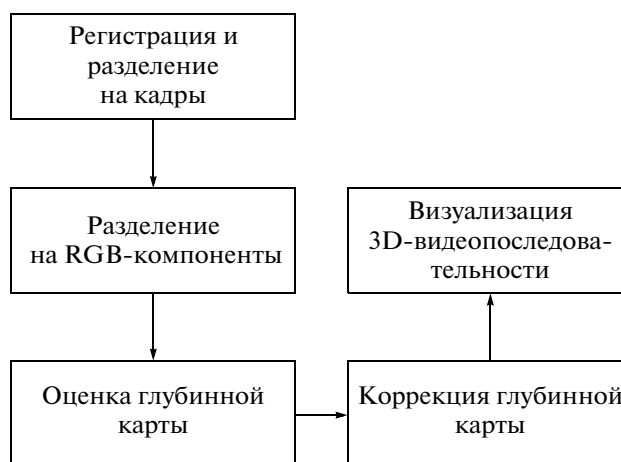


Рис. 1. Блок-схема предлагаемого алгоритма.

Институт радиотехники и электроники
им. В.А. Котельникова
Российской Академии наук, Москва
Национальный политехнический институт Мексики,
Мехико

Научно-технологический центр
уникального приборостроения
Российской Академии наук, Москва

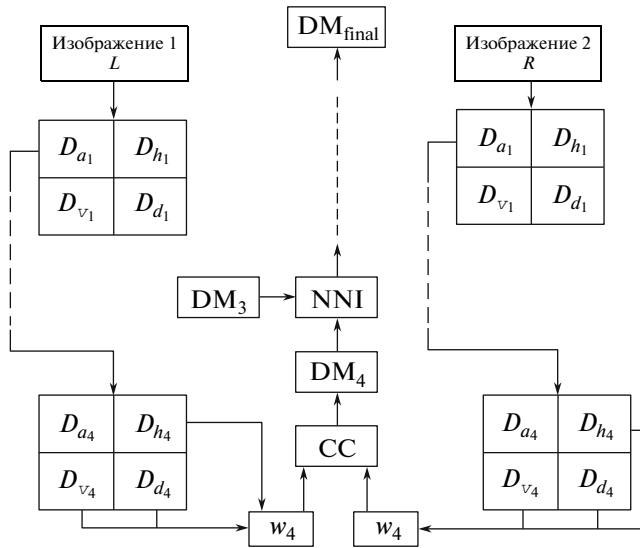


Рис. 2. Блок-диаграмма алгоритма для оценки глубинного изображения.

Оценка глубинной карты на декомпозиционном уровне s (DM_s) для каждой стереопары (L_s и R_s) основана на вычислении нормализованной функции кросс-корреляции (CC_s) в скользящем окне обработки (SO):

$$CC_s(x, y) = \frac{W_{L_s}(i, j)W_{R_s}(x + i, y + j)}{\left(\sum_{i, j \in SO} W_{L_s}^2(i, j) \cdot \sum_{i, j \in SO} W_{R_s}^2(x + i, y + j) \right)^{1/2}}. \quad (3)$$

Алгоритм вычисляет на каждом из четырех уровней декомпозиции для текущей пары соседних кадров в последовательности нормализованные значения вейвлетов W_s , на основании которых затем вычисляются кросс-корреляции CC_s и оценки глубинных карт. Метод NNI используется для интерполяции между текущей и более точной оценками для глубинной карты на каждом уровне декомпозиции при движении от грубой к более точной оценке. Финальная оценка глубинной карты формируется с учетом всех промежуточных ее оценок, как это показано на рис. 2.

Классические методы в построении анаглифов: фотошоп, наименьших квадратов и др. — характеризуются ошибками в виде артефактов и эффектов призраков, а также потерями в качестве цветов. Для улучшения точности оценок глубинных карт и уменьшения указанных ошибок используется метод динамического сжатия в анаглифе [9], состоящий в сжатии глубинного изображения в соответствии с формулой $D_{new} = aD^P$, где D_{new} — новая карта глубины, $0 < a < 1$ и $0 < P < 1$. На завершающем этапе при формировании ана-

глифа используется процедура интерполяции в соответствии с методом NNI.

3. Свойства новой процедуры и известных в литературе алгоритмов 3D-визуализации были исследованы на основе стандартных численных критериев: количество “плохих” пикселей (QBP) [6] и оценка структурного индекса подобия (SSIM) [10], а также с использованием визуального субъективного анализа глубинных карт и генерированных анаглифов.

Значения критерия QBP вычисляются по следующей формуле для каждого из исследуемых изображений:

$$QBD = \frac{1}{N} \sum_{x, y} |d_E(x, y) - d_G(x, y)|^2, \quad (4)$$

где N — число пикселей в изображении или кадре, d_E и d_G — оцененные и истинные карты глубин соответственно.

Величины SSIM-критерия определяются так:

$$SSIM(x, y) = [l(x, y)] \cdot [c(x, y)] \cdot [s(x, y)], \quad (5)$$

где функции люминесцентного $l(x, y)$, контрастного $c(x, y)$ и структурного подобия $s(x, y)$ вычисляются следующим образом:

$$l(x, y) = \frac{2\mu_X(x, y)\mu_Y(x, y) + C_1}{\mu_X^2(x, y) + \mu_Y^2(x, y) + C_1},$$

$$c(x, y) = \frac{2\sigma_X(x, y)\sigma_Y(x, y) + C_2}{\sigma_X^2(x, y) + \sigma_Y^2(x, y) + C_2}, \quad (6)$$

$$s(x, y) = \frac{\sigma_{XY}(x, y) + C_3}{\sigma_X(x, y)\sigma_Y(x, y) + C_3}.$$

Здесь X — оцениваемое изображение, Y — истинное изображение, μ и σ определяют среднее значение и среднеквадратическое значение соответственно для X - или Y -изображений и $C_1 = C_2 = C_3 = 1$.

Были исследованы тестовые синтетические изображения Aloe, Wood1, Lampshade1 и Venus (размер 370×433 пикселей; <http://vision.middlebury.edu/stereo/data>). Дополнительно тестовые цветные видеопоследовательности в формате Avi Coastguard и Flowers (обе 300 кадров, размер 480×720 пикселей, <http://trace.eas.asu.edu/yuv/index.html>) и отснятая реальная видеопоследовательность Video Test (200 кадров, размер 480×720 пикселей) использовались при генерировании трехмерных видеопоследовательностей.

При формировании глубинных изображений кросс корреляция CC вычислялась в скользящем окне (SO) 5×5 пикселей, динамическое сжатие применялось для всех исследуемых алгоритмов с параметрами $a = P = 0.5$. Метод L&K был реализован в соответствии с работой [3], SSD-алгоритм — в соответствии с [4], GEEMSF-процедура повторяла предложенное в оригинальной работе [5]. В новом алгоритме, основанном на вейвлет-пре-

Таблица 1. Значения критериев QBP и SSID для синтетических стереопар изображений

Изображение/критерий	L&K	SSD	GEEMSF	WF Bio6.8	WF Coifl	WF Haar	WAF π_6	M-WF Coifl	M-WAF π_6
Aloe/SSIM	0.3983	0.6166	0.3017	0.9267	0.5826	0.5776	0.9232	0.5826	0.9232
Aloe/QBP	0.1121	0.4722	0.9190	0.0297	0.4517	0.4420	0.0130	0.4490	0.0111
Wood1/SSIM	0.1089	0.7142	0.7051	0.9367	0.7096	0.7072	0.9448	0.7096	0.9448
Wood1/QBP	0.1316	0.2376	0.2100	0.1258	0.2400	0.2402	0.1180	0.2400	0.0919
Lampshade1/SSIM	0.0861	0.6320	0.3124	0.7061	0.7061	0.7081	0.6897	0.7061	0.7619
Lampshade1/QBP	0.2430	0.2800	0.3410	0.2072	0.2071	0.2071	0.2017	0.2071	0.1426
Venus/SSIM	0.1990	0.4320	0.2145	0.5979	0.4530	0.4472	0.4604	0.4530	0.6947
Venus/QBP	0.3084	0.1428	0.2013	0.1694	0.5014	0.5010	0.1930	0.5011	0.1091

образовании, использованы как классические типы вейвлет-функций: койфлет (Coiflets), добеши (Daubechies), биортогональные (Biorthogonal) и симплет (Symlets), так и вейвлеты, основанные на атомарных функциях (up, fur и π).

Как видно из табл. 1, оптимальные значения для критериев QBP и SSIM достигаются при оценке глубинных карт для анализируемых изображений в случае применения алгоритма, основанного на M-WAF π_6 . Следующий по качеству алгоритм использует WAF π_6 . Рисунок 3 демонстрирует лучшее восстановление глубинных карт в случае применения алгоритма, основанного WAF, по сравнению с другими известными методами. Заключительный этап 3D-визуализации был выполнен с использованием техники формирования анаглифа по двум соседним изображениям 2D-видеопоследовательности [9].

Ряд перспективных алгоритмов был реализован на процессоре цифровой обработки EVM DM642 фирмы "Texas Instruments" [11], что позволяет генерировать 3D-видеопоследовательности прак-

тически в режиме реального времени. Здесь использовано последовательное соединение трех процессоров DSP EVM DM642, где первый и второй процессоры реализуют регистрацию текущих кадров 2D-последовательности и формирование глубинных карт, а третий DSP позволяет сгенерировать анаглиф и его визуализацию на дисплее. Такая конфигурация цифровых процессоров существенно увеличила скорость обработки, доведя ее до 20 (240 × 360 пикселей, M-WAF) и 30 цветных кадров в секунду (240 × 360 пикселей, WAF), что соответствует скорости фильма (табл. 2). Традиционные алгоритмы L&K и SSD, которые показали худшее качество в восстановлении трехмерных видеопоследовательностей, требуют для визуализации 3D-кадра в MatLab соответственно 22.59 и 16.26 с, что показывает их неперспективность для обработки в реальном времени.

Измеренные величины времени обработки включали в себя период от регистрации 2D-видеопоследовательности до визуализации сгенерированного анаглифа на мониторе, что позволяет с

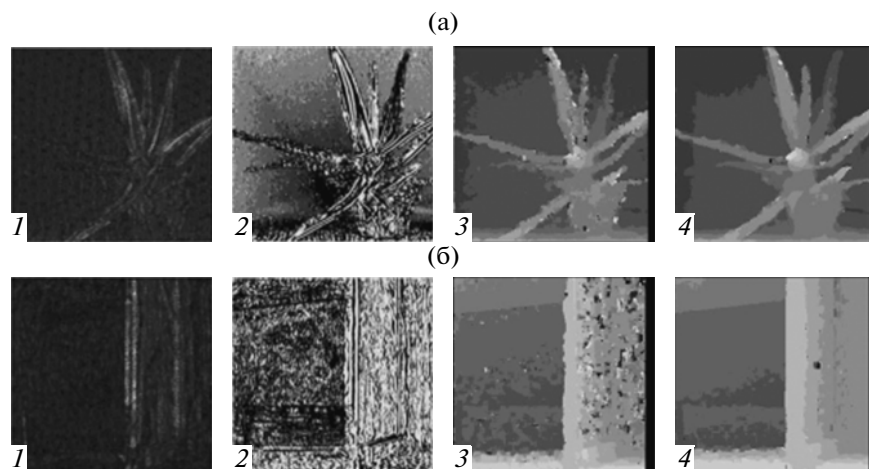


Рис. 3. Глубинные карты, полученные для синтетических изображений Aloe (а) и Wood1 (б) на основе алгоритмов L&K (1), SSD (2), WF BIO 6.8 (3), M-WAF π_6 (4).

Таблица 2. Время, требуемое для различных алгоритмов, при обработке видеопоследовательностей (время/кадр, с)

Алгоритм	Matlab, 480 × 720 пикселей	Три последовательных DSP	
		240 × 360 пикселей	480 × 720 пикселей
Классические вейвлеты Coif1, Db6.8, Haar	6.16	0.031	0.071
WAF (ур, fir_4 , π_6)	6.19	0.031	0.071
М-классические вейвлеты Coif1, Db6.8, Haar	6.76	0.048	0.08
М-WAF (ур, fir_4 , π_6)	6.77	0.049	0.08

помощью линз с фильтрами синего и красного света для левого и правого глаз соответственно наблюдать 3D-видеопоследовательность.

СПИСОК ЛИТЕРАТУРЫ

1. *Ideses I., Yaroslavsky L.* // Lect. Notes. Comp. Sci. 2004. V. 3212. P. 273–280.
2. *Smolic A., Kauff P., Knorr S., et.al.* // Proc. IEEE. 2011. V. 11. № 4. P. 607–625.
3. *Fleet D.J.* Measurement of Image Velocity. Massachusetts: Kluwer Acad. Publ., 1992.
4. *Beauchemin S.S., Barron J.L.* // ACM Comput. Surv. 1995. V. 27. № 3. P. 433–466.
5. *Alagoz B.B.* // OncuBilim Algorithm and Syst. Labs. 2008. V. 8. № 4. P. 1–12.
6. *Bhatti A., Nahavandi S.* // Stereo Vision. 2008. Ch. 2. P. 27–48.
7. *Gulyaev Yu.V., Kravchenko V.F., Pustovoi V.I.* // Doklady Mathematics. 2007. V. 75. № 2. P. 325–332. // ДАН. 2007. Т. 413. № 3. С. 320–328.
8. *Kravchenko V., Perez-Meana H., Ponomaryov V.* Adaptive Digital Processing of Multidimensional Signals with Applications. Moscow: Fizmatlit, 2009.
9. *Ideses I., Yaroslavsky L.* // J. Opt. A: Pure and App. Opt. 2005. V. 7. P. 755–762.
10. *Malpica W.S., Bovik A.C.* // Proc. IEEE Int. Conf. on Acoustics, Speech and Sign. Proc. 2009. P. 1149–1152.
11. Texas Instruments. MS320DM642 Evaluation Module with TVP Video Encoders. Technical Reference. 507345-0001. Texas Instruments. 2004.