

УДК 004.93

О.Л. Вовк, к.т.н., доц.,  
Донецкий национальный технический университет, г. Донецк, Украина  
vovk.olga@gmail.com

## Модифицированный метод k-средних для выделения кластеров кадров

*В работе предложена модификация статистического метода k-средних для выделения регионов кадров видеопоследовательности. Рассмотрена возможность использования предложенного метода для поиска видеоданных.*

**Ключевые слова:** метод k-средних, кластеризация, кадр, видеопоследовательность, поиск видеоданных, цветовое пространство LUV

### Введение

Создание эффективных (с точки зрения, быстродействия и точности) методов поиска видеоданных стало одной из первостепенных задач с момента появления цифрового видео.

Задача распознавания видеоданных наиболее часто формулируется в двух следующих запросах [1,2]:

- найти видеофильмы, имеющие ключевой стоп-кадр, похожий на образец запроса;
- упорядочить видеоматериалы коллекции по их подобию заданному видео образцу.

Эти запросы возникают в системах видеонаблюдения, анализа видеоданных, тематической классификации видеоинформации. Для корректного удовлетворения таких запросов видеоданные предварительно разбиваются на сегменты (отдельные фрагменты) видеоинформации.

Целью работы является разработка модифицированного метода k-средних для выделения кластеров кадров видеоданных. Характеристики полученных кластеров могут быть использованы для содержательного поиска в видеоколлекциях.

Задачи исследования: выбор цветового пространства для обработки изображений, усовершенствование стандартного метода k-средних для выделения кластеров кадров, анализ быстродействия полученного метода, применение разработанного метода для поиска видеоданных.

### Выбор цветового пространства

Для получения значимых результатов при решении задачи выделения кластеров кадров, необходимо начать с выбора цветового пространства, которое бы с максимальной точностью отражало специфику поставленной задачи.

Равномерное цветовое пространство CIE LUV, известное также как CIE 1976 ( $L^*$ ,  $U^*$ ,  $V^*$ ) – пространство цветов, введенное Международной комиссией по освещению (МКО) в 1976 году для того, чтобы минимизировать искажение расстояния между цветовыми характеристиками [3].

В пространстве цветов CIE LUV значение светлоты отделено от значения хроматической составляющей цвета. Светлота задана координатой  $L$  (изменяется от 0 до 100, то есть от самого темного до самого светлого), хроматическая составляющая — двумя декартовыми координатами  $u$  и  $v$ . Первая обозначает положение цвета в диапазоне от зеленого к красному (при увеличении), вторая — от синего к фиолетовому [4, 5]. Параметры  $u$  и  $v$  изменяются в диапазоне от -100 до 100. Если  $u$  и  $v$  равны нулю, то, изменяя  $L$ , получаем цвета, которые являются градациями серого.

Главным преимуществом данного цветового пространства является то, что оно разрабатывалось специально для численного измерения разницы между цветами. Для измерения цветовой разницы между двумя пикселями можно использовать Евклидово расстояние:

$$\Delta = \sqrt{(L_1 - L_2)^2 + (u_1 - u_2)^2 + (v_1 - v_2)^2}, \quad (1)$$

где  $(L_1, u_1, v_1)$  – параметры пикселя первого цвета,  $(L_2, u_2, v_2)$  - параметры пикселя второго цвета.

При расстоянии между двумя цветами  $\Delta > 5$  большинство людей уже замечают разницу, при  $\Delta > 10$  разница между цветами заметна всем. В этом и состоит основное преимущество этого пространства. Оно учитывает восприятие цветов человеком, и разница между цветами определяется очень простой формулой. Стоит отметить, что разница цветности  $\Delta E_{uv} = 13$  соответствует разности светлоты  $\Delta L = 1$ :

$$\Delta E_{uv} = \sqrt{(u_1 - u_2)^2 + (v_1 - v_2)^2}, \quad (2)$$

$$\Delta L = \sqrt{(L_1 - L_2)^2}.$$

Задание показателя светлости  $L$  отдельным параметром позволяет наиболее точно определить расстояние между цветовыми характеристиками пикселей, близких по цветности, но разных по светлости. Так как в отличие от стандартных цветовых пространств RGB, CMYK, HLS, где при изменении светлоты цвета меняются все параметры, в CIE LUV изменяется значение только параметра  $L$ . Это позволяет наиболее точно определить расстояние между двумя близкими по цветности, но разными по светлоте цветами.

Для перехода от цветовой модели RGB к цветовой модели CIE LUV сначала надо выполнить переход в цветовое пространство CIE XYZ.

Цветовая модель CIE XYZ построена на основе зрительных возможностей так называемого «Стандартного наблюдателя», т.е. гипотетического наблюдателя, возможности которого были тщательно изучены и зафиксированы в ходе проведенных комитетом CIE длительных исследований человеческого зрения [5].

Существенный минус цветовой модели CIE XYZ – это то, что она перцептивно (визуально) равномерна и не может быть использована для вычисления расстояний между цветами. Переход от пространства RGB к пространству CIE XYZ осуществляется следующим образом [6]. Сначала параметры  $R, G, B$  нормализуются:

$$\begin{pmatrix} R^* \\ G^* \\ B^* \end{pmatrix} = \begin{pmatrix} R/255 \\ G/255 \\ B/255 \end{pmatrix}. \quad (3)$$

Далее осуществляем переход в пространство CIE XYZ:

$$\begin{pmatrix} X \\ Y \\ Z \end{pmatrix} = \begin{pmatrix} 0.412453 & 0.35758 & 0.180423 \\ 0.212671 & 0.71616 & 0.072169 \\ 0.019334 & 0.119193 & 0.950227 \end{pmatrix} \times \begin{pmatrix} R^* \\ G^* \\ B^* \end{pmatrix}. \quad (4)$$

Переход между пространствами CIE XYZ и CIE LUV можно описать таким образом:

$$L = \begin{cases} 116 \times \sqrt[3]{\frac{Y}{Y_n} - 16}, & \frac{Y}{Y_n} > \left(\frac{6}{29}\right)^3, \\ 903.3 \times \frac{Y}{Y_n}, & \frac{Y}{Y_n} \leq \left(\frac{6}{29}\right)^3, \end{cases} \quad (5)$$

$$u' = \frac{4X}{X + 15Y + 3Z} = \frac{4x}{-2x + 12y + 3},$$

$$u = 13L(u' - u_n),$$

$$v = 13L(v' - v_n).$$

Для определения параметров  $Y_n, u_n$  и  $v_n$ , вводится понятие белой точки. Белая точка – это пара параметров цветности  $(x, y)$ , которые

определяет эталон белого цвета для разных источников света. Международная комиссия по освещению (МКО) составила таблицу белых точек для источников цвета разной яркости. При этом значение компоненты  $Y$  белой точки в XYZ нормализовано до 100 (в приведенных выше формулах (5)  $Y_n$  как раз соответствует нормализованной  $Y$  компоненте). Параметры  $u_n$  и  $v_n$  вычисляются по тем же формулам, что  $u'$  и  $v'$ , в которых используются координаты  $x$  и  $y$  белой точки.

Итак, все расчеты будем вести в цветовой модели CIE LUV, которая максимально адаптирована к человеческому зрению.

### Модифицированный метод $k$ -средний для выделения кластеров кадров

Рассмотрим метод выделения кластеров кадров, который основан на статистическом методе кластеризации  $k$ -средних [7].

Отличительной особенностью предложенного метода является способ нахождения начального числа кластеров.

Введем обозначения. Имеем изображение  $f^0$ , которое состоит из  $k^0$  пикселей  $p_{sg}^0, s \in [1, w^0], g \in [1, h^0]$ . Количество пикселей  $k^0 = w^0 * h^0$ , где  $w^0$  – ширина изображения, а  $h^0$  – его высота. Каждый пиксель  $p_{sg}^0$  задается тремя параметрами  $p_{sg}^0 = \{L_{gs}, u_{gs}, v_{gs}\}$ , которые соответствуют светлоте, переходу от зеленого к красному и переходу от синего к фиолетовому соответственно. Таким образом, изображение  $f^0$  можно записать в следующем виде:

$$f^0 = \{p_{sg}^0 = \{L_{gs}, u_{gs}, v_{gs}\} | s \in [1, w^0], g \in [1, h^0]\}. \quad (6)$$

Современные цифровые изображения и видеофайлы имеют достаточно высокую разрешающую способность, поэтому в работе [О12] было предложено группировать пиксели множества  $P^0 = \{P_{11}^0, \dots, P_{wh}^0\}$  в группы по 16 пикселей, размером  $4*4$ .

Ширина нового изображения  $w = w/4$ , высота  $h = h/4$ . Каждый пиксель  $p_{xy}$  задается тремя параметрами  $(L_{xy}, u_{xy}, v_{xy})$ , которые рассчитываются следующим образом:

$$L_{xy} = \frac{\sum_{g=yn}^{yk} \sum_{s=xn}^{sk} L_{sg}}{16},$$

$$u_{xy} = \frac{\sum_{g=yn}^{yk} \sum_{s=xn}^{sk} u_{sg}}{16}, \quad (7)$$

$$v_{xy} = \frac{\sum_{g=yn}^{yk} \sum_{s=xn}^{sk} v_{sg}}{16}.$$

При  $xn = 4(x-1)+1, xk = 4(x-1)+4, yn = 4(y-1)+1, yk = 4(y-1)+4$ .

Перейдем к определению начального числа кластеров.

Для начального распределения множества пикселей  $P$  по кластерам необходимо знать параметр  $Cl$ , который соответствует количеству кластеров изображения  $f$ .

На начальном этапе  $Cl$  можно определить с помощью метода глобальных гистограмм.

Каждый пиксель  $p_{xy}$  изображения  $f$  задается тремя параметрами  $(L_{xy}, u_{xy}, v_{xy})$ , значения которых лежат в следующих диапазонах:

$$L_{xy} \in [0,100], u_{xy} \in [-100,100], v_{xy} \in [-100,100]. \quad (8)$$

Все пространство цветов представляется в виде трехмерного куба, каждая ось которого соответствует одному из трех параметров  $L_{xy}, u_{xy}, v_{xy}$ . Если каждую ось куба разбить на 4 равные части, то получим 64 равных прямоугольных параллелепипеда. Каждый такой параллелепипед представим в виде множества цветов  $V_g$  ( $g=1..64$ ), компоненты  $(L, u, v)$  которых попали в  $g$ -ый параллелепипед. Для кадра  $f$  задается множество  $H$ , каждый элемент  $h_g$  которого отображает часть каждого подмножества цветов  $V_g$  в общей цветовой гамме изображения:

$$h_g = \frac{N_g}{\sum N_g}, \quad (9)$$

где  $N_g$  – количество точек, которые попали во множество  $V_g$ .

Рассчитаем среднее значение множества  $H$ :

$$h^c = \frac{\sum_{g=0}^{64} h_g}{\sum_{g=0}^{64} [h_g \neq 0]}. \quad (10)$$

Начальное значение числа кластеров  $Cl$  равно числу элементов  $h_g$ , которые превышают среднее значение:  $Cl = \sum_{g=0}^{64} [h_g > h^c]$ .

Перейдем к рассмотрению основных этапов метода.

Первый этап метода. Изначально множество пикселей  $P$  изображения  $f$  делится на  $t$  подмножеств  $t=1..Cl$ . Каждое  $t$ -ое подмножество – это  $t$ -ый кластер.

Дальше для каждого кластера  $y$  вычисляется положение центра кластера  $p_y^c$ , которое есть средним значением всех пикселей, входящих в этот кластер.

Второй этап метода представляет собой итеративный процесс, который продолжается до тех пор, пока множество номеров кластеров  $Y$  не перестает изменяться, т.е. все пиксели  $p_i$ , которые принадлежат кластеру  $y$  на итерации  $q$ , принадлежат этому же кластеру и на итерации  $(q+1)$ :

$$Y_q = Y_{q+1}. \quad (11)$$

Каждому пикселю  $p_i$  снова задается номер кластера  $y$ , такой, что расстояние от пикселя  $p_i$  до центра кластера минимально:

$$y_i = \arg \min_{y \in Y} \rho(p_i, p_y^c). \quad (12)$$

Как мера близости используется Евклидово расстояние:

$$\rho(p_i, p_y^c) = \sqrt{(L_i - L_y^c)^2 + (u_i - u_y^c)^2 + (v_i - v_y^c)^2}. \quad (13)$$

Затем пересчитываются значения центров кластеров. Если условие (11) выполнено, то все пиксели распределены по кластерам.

Третий этап метода представляет собой итеративный процесс, который повторяет первый и второй этапы. Условием завершения третьего этапа является превышение среднего межкластерного расстояния  $Q^c$  заданного количества  $Q^0$ .

$$Q^c = \frac{\sum_{t=1}^c \frac{\sum_{i=1}^{n_c} [y_i = t] \rho(p_i, p_t^c)}{n_c}}{Cl}, \quad (14)$$

где  $n_c$  – количество пикселей, которые попали в кластер с номером  $t$ .

При удовлетворении условия  $Q^c > Q^0$  – третий этап метода завершается.

$Q^0$  подбирается экспериментально, к примеру, для базы изображений исследователя Ванга [8] было получено  $Q^0=0,078$ .

### Анализ быстродействия методов кластеризации

Для проведения экспериментов было отобрано шесть видеофайлов, каждый из которых состоит примерно из 5000 кадров. Характеристики файлов приведены в таблице 1.

Таблица 1 – Характеристики видеофайлов

Номер файла	Количество кадров	Разрешение, пикселей
1	4798	320*224
2	5255	320*192
3	6107	320*208
4	7908	320*240
5	5523	320*240
6	5010	320*208

Практические эксперименты по кластеризации кадров видеофайлов проводились на персональном компьютере с процессором Intel(R) Core2 Duo T5500 @1,66 GHz, и объемом оперативной памяти 2,5 Гб.

Результаты сравнения скорости выделения кластеров кадров стандартным методом  $k$ -средних и модифицированным методом  $k$ -средних приведены на рис. 1.



Рисунок 1 – Быстродействие выделения кластеров кадров рассматриваемыми методами

Полученные результаты показали преимущество разработанного метода для изображений с высокой энтропией (большим числом кластеров).

### Применение модифицированного метода k-средних для поиска видеоданных

Математическую модель поиска видеоданных по изображению-образцу (шаблону) можно представить следующим образом.

Имеем базу изображений  $T$ , которая состоит из  $n$  ключевых кадров видеофайлов  $f_i$ , т.е.  $T = \{f_1, f_2, \dots, f_n\}$ . Кадр  $f_i$  имеет размер  $w_i \times h_i$  пикселей. Каждому пикселю  $p_{ixy}$  ключевого кадра  $f_i$  соответствует три цветовых компонента  $p_{ixy} = \{L_{ixy}, u_{ixy}, v_{ixy}\}$ . Математически ключевой кадр можно записать в виде:

$$f_i = \{p_{ixy} = \{L_{ixy}, u_{ixy}, v_{ixy}\} \mid x \in [1, w_i], y \in [1, h_i]\}, \quad (15)$$

Кроме базы ключевых кадров на вход системы контекстного поиска подается изображение-образец поиска  $f^0$ , заданное в виде (15).

Каждый ключевой кадр  $f_i$  базы данных  $T$  соответствует определенному видеофайлу. Т.е. поиск видеофайла по изображению-образцу  $f^0$  сводится к поиску соответствующего ключевого кадра  $f_i$ , поиск которого основан на сравнении характеристик выделенных кластеров ключевого кадра и изображения, заданного в качестве образца. Способ выделения кластеров кадров подробно описан выше. Процедура выделения кластеров для изображения-образца аналогична процедуре выделения кластеров для ключевого кадра.

Каждому номеру кластера  $u$  ключевого кадра  $f_i$  соответствует вектор характеристик

$C_y = \{c_{y1}, c_{y2}, \dots, c_{yM}\}$ , а каждому кластеру  $o$  изображения образца соответствует вектор характеристик  $C^0 = \{c_1^0, c_2^0, \dots, c_M^0\}$ , где  $M$  – количество характеристик кластера.

При вычисления вектора характеристик  $C_y$  кадра  $f_i$  (или  $C^0$  образца  $f^0$ ) учитываются характеристики цвета и геометрические характеристики местоположения пикселей  $p_{ixy}$  (или  $p_{xy}^0$ ).

В основе сравнения кадров и изображения-образца поиска по характеристикам – Евклидово расстояние ( $D = \{d_1, d_2, \dots, d_n\}$ ), где  $d_i$  – расстояние от кадра  $f_i$  до образца  $f^0$ .

Множество расстояний  $D$  упорядочивается по возрастанию. Результатом поиска есть массив  $R$  ключевых кадров видеофайлов,  $R = \{f_i \mid i \in L\}$ , где  $L$  – заданное пользователем количество результатов поиска. Каждому кадру  $f_i$  базы данных  $T$  соответствует расстояние  $d_i$  множества  $D$ .

### Заключение

Предложенный метод кластеризации позволяет уменьшить количество перераспределений пикселей по кластерам за счет определения начального (стартового) числа кластеров. Метод определения начального числа кластеров основан на методе глобальных гистограмм. Переход от стандартного пространства цветов RGB в пространство цветов CIE LUV позволяет более точно определять расстояние между цветовыми характеристиками.

Научная новизна работы заключается в модификации статистического метода кластеризации k-средних путем добавления этапа предварительного определения числа кластеров кадров.

Практическая значимость работы состоит в разработке системы кластеризации кадров видеоданных, на основе которой в дальнейшем планируется разработка системы контекстного поиска видеoinформации.

Дальнейшие исследования будут направлены на разработку методов расчета характеристик кластеров, которые позволят обеспечить оптимальное сочетание скорости и точности содержательного поиска кадров в коллекциях видеофайлов.

**Список использованной литературы**

1. Байгарова Н.С., Бухштаб Ю.А. Проект «Кинолетопись России»: представление и поиск видеoinформации / Н.С. Байгарова, Ю.А. Бухштаб // I Всероссийская конференция "Электронные библиотеки". – Санкт-Петербург, 1999. – с.209-215.
2. Loui A. C., Gatica-Perez D. Video structuring by probabilistic merging of video segments / A. C. Loui, D. Gatica-Perez.. – USA:US\_Patent7, 2007. – 17 p.
3. Chen C.H., Pau L.F., Wang P.S.P. The Handbook of Pattern Recognition and Computer Vision (2<sup>nd</sup> Edition)/ C.H. Chen, L.F. Pau, P.S.P. Wang. – World Scientific Publishing Co., 1998. – 1004 p.
4. Fairchild M. D. Color Appearance Models. Reading / M. D. Fairchild. - MA: Addison-Wesley, 1998. – 417 p.
5. Blesslin Elizabeth C.P., Kingsly Devi K. U. Spectral Clustering of Images in LUV Color Space by Spatial-Color Pixel Classification / C.P. Blesslin Elizabeth, K. U. Kingsly Devi. // International Journal of Computer Applications (0975 – 8887). – 2010. – vol. 3, №9. – P.2-5.
6. Куликов А.И., Овчинникова Т.Э. Алгоритмические основы современной компьютерной графики / А.И. Куликов, Т.Э. Овчинникова. – 2004. – [Электронный ресурс] – Режим доступа: <http://www.intuit.ru/department/graphics/graphalg/2/5.html>
7. Wang J. Z., Li J., Wiederhold G. SIMPLiCity: Semantics-Sensitive Integrated Matching for Picture Libraries / J. Z. Wang, J. Li, G. Wiederhold // IEEE Transactions on Pattern Analysis and Machine Intelligence. – 2001. – vol. 23, №9. – P. 947-963.
8. База данных изображений группы исследователей под руководством Wang P.S.P [Электронный ресурс] – Режим доступа: <http://wang.ist.psu.edu/~jwang/test1.tar>

Надійшла до редколегії 30.01.2011

**О.Л. Вовк**

Донецкий національний технічний університет

**O.L. Vovk**

Donetsk National Technical University

**МОДИФІКОВАНИЙ МЕТОД К-СЕРЕДНІХ  
ДЛЯ ВИДІЛЕННЯ КЛАСТЕРІВ КАДРІВ.****A MODIFIED K-MEANS METHOD OF  
CLUSTERING TO EXTRACT CLUSTERS OF  
THE FRAMES.**

У роботі запропоновано модифікацію статистичного методу к-середніх для виділення регіонів кадрів відеопослідовності. Розглянуто можливість використання запропонованого методу для пошуку відеоданих.

Modification of k-means statistical method for the extract of regions frames of videosequence is offered in this work. Possibility of the use of the offered method for the search of videoinformation is considered.

**Ключові слова:** метод к-середніх, кластеризація, кадр, відеопослідовність, пошук відеоданих, кольорний простір LUV

**Keywords:** k-means method, clusterization, frame, videosequence, search of videoinformation, LUV colour space