

УДК 004.032.26

# Нейронный классификатор смены сцены и шумоподобности кадров видеопоследовательности

© Автор, 2009

Ю.В. Слынько -

аспирант Московского Физико-Технического Института (Государственного Университета), кафедры Информационных систем при ОАО «МАК «Вымпел», E-mail: [UrasSL@yandex.ru](mailto:UrasSL@yandex.ru)

Представлен метод детектирования смены сцены, шумоподобности кадров видеопоследовательности. На основе использования метода метрик и нейросетевого подхода предложено решение задачи классификации. Показано, что метрики являются инвариантными по отношению к сдвигу.

**Ключевые слова:** нейронный классификатор, инвариант к сдвигу, видеопоследовательность.

## Введение и постановка задачи

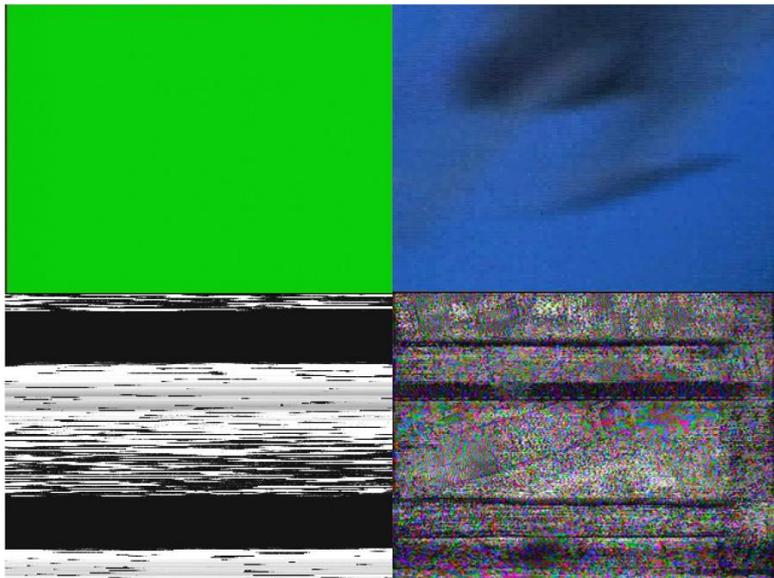
В данной работе рассматривается задача определения качества видеопоследовательности. Здесь под качественной последовательностью подразумевается такая последовательность, для которой возможна дальнейшая обработка. В качестве обработки рассматривались, прежде всего, алгоритмы стабилизации и сопровождения объектов [1, 2]. Если алгоритм стабилизации определяет параметры геометрических преобразований неправильно, то выходное изображение, «стабилизированное» в соответствии с такими неправильными параметрами, «скачет» больше, чем входное, поэтому следует отсеивать кадры, обработка которых не может быть выполнена достаточно качественно.

Суть проблемы в том, что алгоритм стабилизации оценивает параметры геометрических преобразований кадров путем нахождения минимума некоторой целевой функции. Таким образом, он не может работать при несоответствии входных изображений модели геометрических преобразований принятых при разработке алгоритма, а также при слабой обусловленности задачи — когда целевая функция слабо зависит от параметров преобразований.

На практике такие ситуации возможны в случаях (рис. 1): смена сцены; отсутствие полезного сигнала (при наличии или отсутствии шума); высокий уровень помех; некоторые другие случаи, например, съемка с большим увеличением при быстром перемещении камеры, приводит к сильным «смазам» изображения и слишком большим смещениям.

Несмотря на изначальную нацеленность алгоритма на работу с алгоритмами стабилизации и сопровождения объектов, он может быть использован и для более общих задач - определения смены сцены, шумовых кадров и т.д.

В данной области была проделана большая работа разными кол-



**Рис. 1.** Примеры кадров низкого качества. Слева сверху – отсутствие полезного сигнала, справа сверху – съемка с движущейся камеры с большим увеличением, слева внизу – помехи в канале связи при отсутствии сигнала, справа внизу – помехи в канале связи на фоне полезного сигнала

лективами. Особенно много исследований коснулось проблемы поиска смены сцены как, пожалуй, самой востребованной задачи из вышеназванных. Большинство алгоритмов легко снести к простой схеме - вычисление метрик с последующей их классификацией. В простейшем случае - сравнение с порогом. Особую популярность получили алгоритмы, использующие сжатые алгоритмом MPEG-4 последовательности [9 - 12], где в качестве метрик взяты несколько первых коэффициентов разложения в ряд по косинусам. Для получения метрик могут быть также использованы пространственно-временные разложения [6] или разложения по вейвлетам [3]. Существует ряд алгоритмов [5], которые используют сложные алгоритмы анализа сцены, включающие в себя сегментацию и сопровождение объектов, для нахождения смены сцены. В работах [4, 7, 8] представлены методы, основанные на простых метриках, таких как разность последовательных кадров, скорость изменения яркости и анализ гистограмм. Работа [3] особенно интересна тем, что смена сцены в ней не просто детектируется, но выносится решение о «силе» изменения сцены.

Большинство методов используют эмпирические метрики и классификацию путем сравнения с порогами, что не даёт удовлетворительного результата. В данной работе предложена оригинальная метрика, описывающая характер коррелированности последующих кадров друг с другом, инвариантная по отношению к сдвигу изображения. Применение классификации методом нейронных сетей позволило эффективно решить задачу на основе выбранных метрик.

### Выбранные метрики

Степень похожести двух кадров  $F^1$  и  $F^2$  может быть выражена в виде невязки:

$$X_{(F^1, F^2)}(dx, dy) = \frac{1}{S(M)} \sum_{(i,j) \in M(dx, dy)} (F^1_{ij} - F^2_{i+dx, j+dy})^2. \quad (1)$$

Здесь  $M(dx, dy)$  - область перекрытия кадров для заданных сдвигов  $dx$  и  $dy$ , а  $S(M)$  - площадь этой области. Для ускорения алгоритма удобно вместо оригинальных входных кадров использовать сжатые в  $n$  раз кадры.

Эта величина часто используется в качестве метрики, но, как показывает практика, не является хорошей характеристикой связанности кадров, так как не учитывает шум.

Более точной оценкой характера коррелированности кадра является площадь доверительной области нахождения минимума невязки. Если эта область большая, значит невязка имеет шумовой характер, что справедливо для всех рассматриваемых здесь случаев.

Для повышения надежности оказалось целесообразным использовать также и другие, инвариантные к сдвигу, метрики. Так, в качестве метрик были выбраны следующие величины:  $S_b$  — площадь доверительной области минимума функции невязки;  $C_{min}$  - значение невязки в минимуме;  $N_d$  —  $\chi^2$ -критерий разности гистограмм.

Доверительная область минимума функции (1) выражена формулами:

$$\tilde{D} = \{dx_i, dy_i : X_{(F^1, F^2)}(dx_i, dy_i) < m_i + B \cdot \sigma_i\}, \quad (2)$$

где

$$m_i = \frac{X_{min}}{N_p} \cdot N_i, \quad \sigma_i = \frac{X_{min}}{N_p} \cdot \sqrt{2N_i},$$

$$X_{min} = \min_{(dx, dy) \in D} X_{(F^1, F^2)}(dx, dy), \quad N_p = S(M(dx, dy)), \quad N_i = S(M(dx_i, dy_i)),$$

где  $B$  - порог, определяющий доверительную вероятность ошибки;  $S_b$  - является площадью области  $\tilde{D}$ , а  $C_{min}$  - значение невязки в минимуме.

Критерий  $N_d$  может быть вычислен стандартным образом [13] по области перекрытия кадров, сдвинутых относительно друг друга на величину оценки сдвига  $(dx, dy)$ .

### Классификация на основе метрик

Указанные выше метрики используются для классификации кадров видеопоследовательности на качественные и некачественные. Наиболее эффективным классификатором является нейронная сеть.

В данной работе была использована трехслойная нейронная сеть, обученная на представительной обучающей выборке. Она включала в себя 6 последовательностей хорошего качества, снятых в разных условиях (в комнате, на открытой местности, из движущейся машины, с БЛА, при наличии или отсутствии движения камеры и/или объектов в поле зрения) и 6 последовательностей, признанных «плохими» (рис. 1).

Важно, что выход нейронной сети является дробной величиной от -1 до 1. При этом величины больше нуля соответствуют «хорошим» кадрам, а меньше — «плохим». Исходную дробную величину можно использовать для характеристики качества кадра в дальнейшей обработке (например, как параметр регуляризации при определении геометрических трансформаций).

Для обучения нейронной сети был использован алгоритм обратного распространения. Проверались и более простые способы классификации — разделение областей «хороших» и «плохих» кадров в пространстве признаков трехпараметрическими поверхностями — плоскостью, эллипсоидом или параллелепипедом.

Для достижения лучшего результата использовалась простейшая временная фильтрация по критерию из  $N$  по  $K$ , т.е. кадр признавался плохим, если из  $N$  предыдущих кадров  $K$  оказывались плохими. Этот метод весьма эффективен во всех случаях, кроме смены сцены. А, как показало исследование, детектирование смены сцены является наиболее простой задачей из решаемых.

### Результаты и выводы

Основным критерием качества работы представленных алгоритмов является их кривая эффективности — т.е. зависимость количества кадров, ложно детектированных как плохие, от количества недетектированных плохих кадров. Графики этих кривых приведены на рис. 2. На них представлены результаты для наиболее трудных случаев, когда задача неформализуема и не может быть точно решена даже оператором (например, сильный шум на фоне полезного сигнала).

Для сравнения были также исследованы трехпараметрические пороговые классификаторы.

Область - параллелепипед:  
 $S_b < \alpha, X_m < \beta, H_d < \gamma; \quad (3)$

область - полупространство:  
 $\alpha S_b + \beta X_m + \gamma H_d < 1; \quad (4)$

Пороги  $\alpha, \beta$  и  $\gamma$  искались двумя способами - методом полного перебора и методом спуска (см. рис. 2).

В таблице приведены сводные результаты работы алгоритмов, касающиеся отдельных подзадач общей задачи определения «плохих» видеопоследовательностей.

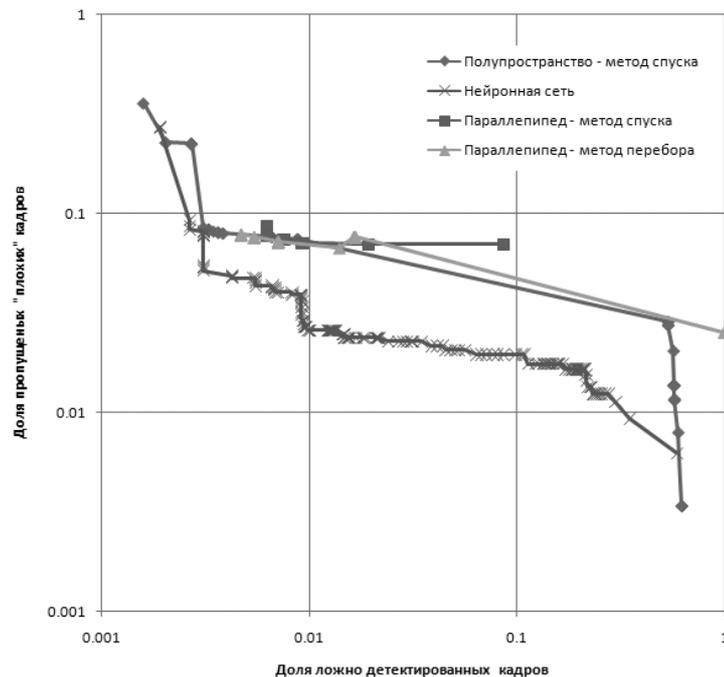


Рис. 2. Кривые эффективности. Для сравнения представлены кривые эффективности, вычисленные для трехпараметрических поверхностей.

Таблица. Результаты работы алгоритмов для определения «плохих» видеопоследовательностей

Скорость работы (кадров в сек.) Pentium M 1700МГц	-160
Вероятность пропуска / вероятность ложного детектирования для задачи детектирования смены сцены	0,2%/<0,01%
Вероятность пропуска / вероятность ложного детектирования для задачи детектирования шумовых кадров при отсутствии полезного сигнала	<0,01%/<0,01%
Вероятность пропуска / вероятность ложного детектирования для задачи детектирования шумовых кадров в присутствии полезного сигнала	<0,01%/0,01%

Таким образом, решена задача оценивания качества входной видео последовательности и детектирование таких событий, как смена сцены, шумоподобность кадра, высокий уровень шумов на фоне полезного сигнала и др.

### Литература

1. Слынько Ю.В., Лагуткин В.Н., Лукьянов А.П. Разработка и исследование робастных алгоритмов реального времени оценивания параметров геометрических преобразований кадров видеопоследовательностей. Радиотехника и электроника. 2007. Т. 52. № 3. С. 351 - 357.
2. Слынько Ю.В. Лукьянов, А.П., Лагуткин В.Н. Гибридный подход к решению задачи оценивания геометрических искажений кадров видеопоследовательностей. // Доклады конф. Цифровая обработка сигналов и ее применение. Вып. VIII - 2. С. 437 - 440.
3. Korpi-Anttila J. Automatic Colour Enhancement and Scene Change Detection of Digital Video. // Graphic Arts in Finland 32(2003)].
4. Ralph M. Ford, Craig Robson, Daniel Temple, and Michael Gerlach. Metrics for Scene Change Detection in Digital Video Sequences. // ICMCS '97, IEEE.
5. Shu-Ching Chen, Mei-Ling Shyu, Cheng-Cui Zhang, R.L. Kashyap. Video scene change detection method using unsupervised segmentation and object tracking, II IEEE International Conference on Multimedia and Expo (ICME). P. 57 - 60. August 22 - 25. 2001. Waseda University, Tokyo, Japan.
6. Chong-Wah Ngo, Ting-Chuen Pong, Hong-Jiang Zhang & Roland T. Chin. Motion-based Video Representation for Scene Change Detection. // International Journal of Computer Vision. Volume 50. Issue 2 (November 2002).
7. Jong Wan Jang and Il Kyun Oh. Performance Evaluation of Scene Change Detection Algorithms. // APCC/OECC '99. V. 2.
8. Xiaoquan Yi and Nam Ling. Fast Pixel-Based Video Scene Change Detection. // ISCAS 2005. V. 4.
9. Georgios Akrivas, Nikolaos. D. Doulamis, Anastasios. D. Doulamis and Stefanos. D. Kollias. Scene Detection Methods for MPEG - Encoded Video Signals. // MELECON 2000. V. 2.
10. Serkan Kiranyaz, Kerem Caglar, Bogdan Cramariuc and Moncef Gabbouj. Unsupervised Scene Change Detection Techniques In Feature Domain Via Clustering And Elimination. // IWDC 2002. Capri, Italy.
11. Edmundo Saez, Jose I. Benavides, Nicolas Guil. Reliable real time scene change detection in MPEG compressed video. // ICME '04. V. 1.
12. Lu H.B., Zhang Y.J., Yao Y.R. Robust Gradual Scene Change Detection. // ICIIP 99. V. 3.
13. Смирнов Н.В. Теория вероятностей и математическая статистика. Избранные труды. М.: Наука. 1970.

Поступила 23 июня 2008 г.

## Neuron Classifier of Noise-Type Frames and Scene Changing in Video

© Author, 2009 Yu. V. Slynko

Method of detection of noise-type frames, scene changes and some other situations in video is proposed. Method is based on metrics, which sufficiently describe character of frames correlation. All proposed metrics are shift-invariant. Neuronet approach is used for solution of classification task.

**Keywords:** neuronal classifier, shift-invariant, in Video.