

МЕТОДЫ ПОСТРОЕНИЯ АЛГОРИТМА КЛАССИФИКАЦИИ И ОЦЕНКИ КАЧЕСТВА ВИДЕОПОСЛЕДОВАТЕЛЬНОСТИ.

Задача классификации ситуаций в видеопоследовательностях весьма популярна в настоящее время. Она востребована в задачах создания автоматизированных баз данных видеопоследовательностей, оценки качества, выделения ключевых кадров при кодировании и т.д.

В данной работе рассмотрены такие ситуации, как смена сцены, шумовые и шумоподобные кадры, кадры с линейной структурой и т.д. Все эти ситуации рассматриваются как некачественное – «плохое» - видео. Таким образом, стоит задача детектирования такого «плохого» видео.

Согласно [1] дуальтернативная задача имеет следующее решение:

$$1\text{-е решение принимается, если } \frac{p_1(\vec{x})}{p_2(\vec{x})} > \frac{p_2(g_{22} - g_{12})}{p_1(g_{11} - g_{21})}, \text{ иначе } - 2\text{-е.} \quad (1)$$

Здесь \vec{x} - вектор наблюдения (минимальная достаточная статистика), $p_i(\vec{x})$ - функция правдоподобия i -ой ситуации, p_i - априорные вероятности ситуаций, g_{ij} - функция потерь.

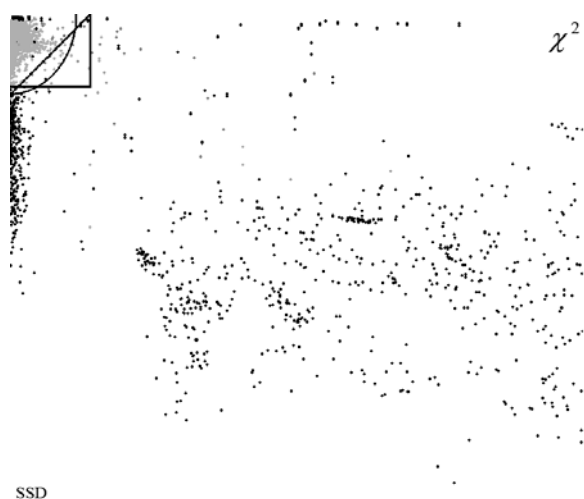


Рис. 1. Проекция трехмерного пространства метрик на метрики χ^2 и SSD. Яркие точки соответствуют «плохим» кадрам. В левом верхнем углу показаны оптимальные трехпараметрические области.

взаимокорреляционной функции последовательных изображений.

В реальности функции правдоподобия в различных ситуациях неизвестны, а размерность задачи очень велика (может достигать до миллионов). Так что необходимо найти способ снижения размерности, т.е. необходимо найти минимальную достаточную (или приблизительно достаточную) статистику. Существует большое количество разнообразных статистик [2], [3]. Исследовав большое количество различных метрик, были выбраны следующие:

- χ^2 -разность гистограмм последовательных изображений
- SSD - Сумма квадратов разностей последовательных изображений
- Sbasin - Размер минимума

На практике задача выбора правильного порога в (1) выглядит труднее задачи выбора достаточных статистик. Фактически, во многих работах предлагается выбирать пороги практически вслепую. Гораздо удобнее построить диаграмму (Рис.1), по осям которой стоят значение выбранных метрик. С использованием такой гистограммы можно достаточно надежно выбрать пороги. Более того, можно точно обозначить область значения метрик, которые характеризуют одну из ситуаций.

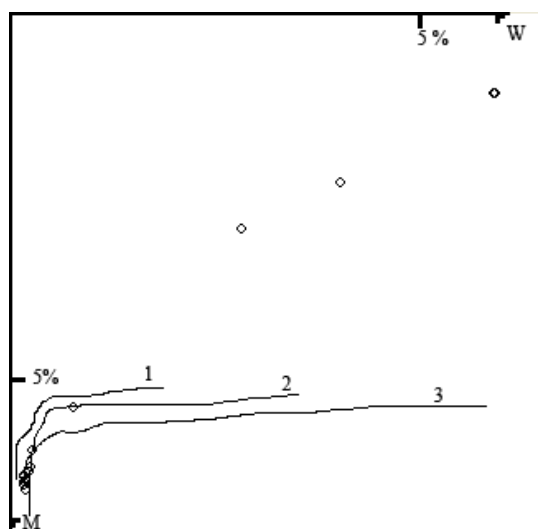


Рис. 2. Кривые эффективности. Линии – рассчитанные перебором (1 – для эллиптической области, 2 – для трех плоскостей вдоль осей, 3 – для одной плоскости). Круги – кривая эффективности, полученная минимизацией критерия (2) для плоскости.

Более эффективный метод построения решающего правила – построение кривой эффективности. Для построения кривой эффективности можно использовать минимизацию следующего критерия:

$$J = W(\vec{a}) + \alpha M(\vec{a}), \text{ где} \quad (2)$$

W – количество «хороших» кадров, определенных как «плохие», M – количество неопределенных «плохих» кадров, α – весовой коэффициент, \vec{a} – вектор параметров порога, характеризующих область «хороших» кадров.

В данной работе исследовались три возможных формы области – ограниченные плоскостью, эллипсоидом или тремя плоскостями, параллельными осям (Рис. 1).

Минимизируя данный критерий по вектору \vec{a} можно получить соотношение ложных и пропущенных детектирований как параметрическую функцию α , и тем самым построить кривую эффективности (Рис. 2).

Однако минимизация данного критерия задача весьма непростая за (счет дискретности и наличия побочных минимумов) и обычные методы спуска, например метод Пауэлла, недействительны. Для улучшения сходимости необходимо вводить регуляризацию, т.е. каждой точке в пространстве метрик приписывать вес, равный 1 вдали от границы области, и плавно сходящийся к 0 на самой этой границе.

Также можно строить кривую эффективности прямым перебором (Рис. 2). Однако, даже задаваясь достаточно мелким шагом, невозможно построить достаточно точную оценку оптимальной кривой эффективности.

Такие параметры обеспечивают детектирование более 99% смен сцены, 99% шумоподобных кадров и не менее 50% шумовых кадров. Данный алгоритм был протестирован на 49 различных видеопоследовательностях и показал немногочисленные сбои в работе только на 3-х из них.

Литература

1. Репин В.Г., Тартаковский Г.П. Статистический синтез. - «Советское радио», 1977.
2. Jong Wan Jang and Il Kyun Oh. // APCC/OECC'99, pp 841-844.

3. Xiaoquan Yi and Nam Ling // ISCAS 2005, pp 3443-3446.