

# Обнаружение и подсчёт числа людей сидящих в креслах во время киносеанса

Альберт Ахриев

Александр Болтнев

June 17, 2005

## 1 Краткий анализ задачи

В рамках проекта CinemaPark нами решалась алгоритмическая часть задачи, состоящей в обнаружении и подсчёте числа людей сидящих в креслах. Сложность задачи обусловлена целым рядом факторов, главным из которых является низкое качество видео. Перечислим эти факторы по порядку

1. Сильный шум, обусловленный наблюдением в темноте;
2. Мерцающий свет от экрана. Причем освещенность колеблется от полной темноты до яркой вспышки;
3. Помехи канала передачи, которые проявляются как лёгкое мерцание с частотой 50 Гц, завал верхних частот, вертикальные полосы, обусловленные ограниченностью спектра аналогового сигнала, и явные помехи в виде горизонтальных полос из-за плохой развязки по питанию;
4. Низкое разрешение кресел дальних рядов и некоторая расфокусировка изображения, связанная с тем, что диафрагма объектива почти полностью открыта (малая глубина резкости), а протяженность зала в глубину достаточно велика;
5. Свет кинопроектора, который отражается от воздушной пыли, и крайне неравномерно и хаотично засвечивает кресла стоящие в зале;
6. Люди сидящие в креслах, в условиях низкого качества видео, порой трудноотличимы от одежды (висящей на спинке кресла или лежащей на сидении), а иногда и от пустого кресла;

7. Люди зачастую принимают необычные позы, а именно, облокачиваются на рядом сидящего соседа так, что спинка кресла становится почти полностью видна, полулежат так, что перекрывают значительную часть спинки соседнего кресла, облокачиваются на спинку впереди стоящего кресла так, что спинка их собственного кресла становится почти полностью видна и т.д.;
8. Люди часто перемещаются, встают, выходят на время, пересаживаются с кресла на кресло, копаются в сумках лежащих на соседних креслах, лежат на переднем ряду и т.п.;
9. Люди имеют очень большой разброс темперамента. Одни сидят часами почти неподвижно, другие мотают головой и руками 50-60% времени;
10. Мы практически не можем разрешить следующую дилемму. Допустим, человек вышел из зала и через некоторое время вернулся, сев на то же или на другое кресло. В другом варианте человек вышел из зала и больше не возвращался, но вошёл другой человек, который, возможно, сел на то же место. В первом случае количество зрителей не изменилось. Во втором случае оно увеличилось на единицу, просто один из зрителей не досидел до конца. Конечно, гипотетически, можно попытаться анализировать яркость одежды, пол и возраст вошедших или вышедших людей, но наши возможности здесь сильно ограничены. Таким образом, существуют трудно разрешимые ситуации приводящие к систематическим ошибкам.

## 2 Подходы к решению задачи

С самого начала мы отказались от методов применяемых в системе *Orwell*, поскольку у нас нет стационарного фона из-за неравномерного и мерцающего освещения. Как следствие, мы отказались (пока) от слежения за человеком, перемещающимся по залу, пересаживающимся с кресла на кресло или покидающим зал. Этот отказ мотивирован ещё и тем, что у нас зал покрывается несколькими камерами, поэтому движущийся объект нужно “передавать” от камеры к камере, а это ведёт к существенному усложнению всей технологии работы. Однако во время практически любого сеанса бывают довольно продолжительные промежутки, когда фон можно считать почти постоянным. Поэтому в дальнейшем, имеет смысл вернуться к задаче прослеживания перемещения отдельных

людей.

Принимая во внимание сроки исполнения работы и наличия серьёзных проблем, перечисленных выше, мы сформулировали следующий подход к решению задачи подсчёта числа людей

1. Алгоритм работает только в области покрывающей примерно 60% спинки кресла и расположенной в верхней её части. Только эта часть кресла видна при любом расположении кресла относительно камеры;
2. Перед началом работы производится съёмка пустого зала, освещённого кинопроектором, в котором не заправлена пленка. Таким образом, мы получаем пустые кресла, освещённые примерно так же, как это происходит во время сеанса. Полученное методом накопления (для снижения шумов) изображение называется *статическим фоном*;
3. Соответствующие кресла на статическом фоне обводятся в специальной программе. В результате мы имеем набор замкнутых кривых, которые очерчивают *зону наблюдения* в пределах каждого кресла, т.е. получается привязка кресел к определённым областям изображения;
4. Во время работы мы сравниваем статический фон с текущим кадром в пределах зон наблюдения. В результате вычисляется значение разности между текущим изображением кресла и его изображением на статическом фоне, полученное из *нормализованной кросс-корреляции*. Значение разности мы называем *статической корреляцией*, хотя это и не совсем математически корректно. Если разность меньше порога, то кресло пустое, иначе кресло занято вещами или человеком. Следует отметить, что нормализованная кросс-корреляция довольно точно отделяет пустые кресла от непустых. Её главное достоинство слабая чувствительность к резким перепадам яркости;
5. Чтобы отличать людей от одежды одной статической корреляции недостаточно. К счастью люди, даже самые флегматичные, достаточно часто сменяют позу. Чтобы обнаруживать движения людей, мы одновременно накапливаем два изображения. Оба изображения складываются в определённой пропорции с предыдущими кадрами. В первом изображении доля предыдущих кадров мала, а текущего велика. Во втором изображении наоборот, доля предыдущих кадров велика, а текущего мала.

Таким образом, второе изображение запаздывает по отношению к первому изображению. Если человек начинает двигаться в кресле, то интегральное значение разности между обоими изображениями становится ненулевой. Эту разность мы называем *динамической корреляцией*, хотя это и не совсем математически корректно. Динамическая корреляция также вычисляется из нормализованной кросс-корреляции, т.е. она слабо-чувствительна к вариациям засветки;

6. Для каждого кресла и подсчитываем два типа интервалов. *Статические интервалы* характеризуют временные промежутки когда статическая корреляция превышала пороговое значение, т.е. кресло было чем-то или кем-то занято. Мы считаем, что кресло может быть занято человеком, если суммарная продолжительность статических интервалов превышает половину продолжительности сеанса;
7. Внутри статических интервалов могут находиться несколько динамических интервалов. *Динамическим интервалом* мы называем временной промежуток, в течение которого значение динамической корреляции превышало пороговое значение. Слишком короткие динамические интервалы трактуются как случайные и не принимаются во внимание, что задаётся в настройках;
8. Если статические интервалы наблюдения данного кресла достаточно продолжительны и объект, занимающий кресло пошевелился несколько раз за сеанс, то мы считаем объект человеком и увеличиваем счётчик;
9. Если суммарная продолжительность динамических интервалов соизмерима с продолжительностью соответствующих статических интервалов, то мы не считаем объект человеком. Последнее условие полезно в следующей ситуации. Допустим, на сидении кресла находится сумка, а само кресло определяется как пустое. Когда человек сидящий на соседнем кресле начинает копошиться в сумке, то он наклоняется так, что ранее пустое кресло теперь определяется как занятое и в тоже время там наблюдается активное движение. Иными словами, период занятости соизмерим с периодом активности, что не является признаком человека сидящего в данном кресле.

В данном описании мы, естественно, опустили ряд важных деталей реализации, однако нами были упомянуты все основные пункты первой версии алгоритма. Более подробное описание с формулами и блок-схемами будет представлено позже.

Несмотря на свою очевидную простоту, алгоритм выдаёт вполне разумные результаты. Очевидный недостаток первой версии чрезмерный эмпиризм в подборе параметров алгоритма. В настоящее время нами практически завершена разработка второй версии алгоритма. Во второй версии мы также проводим анализ только в пределах спинки кресла. Однако теперь наш анализ опирается на более строгий математический подход, включающий в себя процедуру обучения на тестовых видеопоследовательностях. В следующей, третьей версии, мы планируем использовать дополнительный нейросетевой блок распознавания. В значительной мере работа в этом направлении уже проделана.

## References