

M.Turk & A.Pentland: Eigenfaces for recognition, 1991

Анализ статьи

Загонов Дмитрий

Национальный исследовательский университет
"Высшая школа экономики"

18 октября 2025 г.

Актуальность

Цитата

Computational models of face recognition, in particular, are interesting because they can contribute not only to theoretical insights but also to practical applications. Computers that recognize faces could be applied to a wide variety of problems, including criminal identification, security systems, image and film processing, and human-computer interaction.

Цитата

Unfortunately, developing a computational model of face recognition is quite difficult, because faces are complex, multidimensional, and meaningful visual stimuli.

Цитата

Our goal, which we believe we have reached, was to develop a computational model of face recognition that is fast, reasonably simple, and accurate in constrained environments such as an office or a household.

- **Bledsoe (1966)** - один из первых проектов по полуавтоматическому распознаванию лиц. Использовались фидуциальные точки (углы глаз, кончики носа и т.д.), которые оператор вручную отмечал на фотографии.
- **Fischler, Elschlager (1973)** - попытка автоматического измерения похожих признаков при помощи алгоритма линейного встраивания (linear embedding).
- **Kanade (1973)** - глобальная модель ожидаемых характеристик и статистическая классификация по локальным гистограммам яркости (беда с освещением).
- **Midorikawa (1988)** - лица как общие двумерные паттерны, но без структурных свойств (слабая масштабируемость).

Eigenfaces

Идея: получить из изображений лиц их наиболее важные признаки (как глобальные, так и локальные). Признаки при этом не обязательно должны быть интуитивными.

Математически: найти главные компоненты распределения изображений лиц в многомерном пространстве изображений (Sirovich & Kirby, 1987), используя метод PCA (Principal Component Analysis), что позволит снизить размерность пространства признаков.

Определение

Eigenfaces (собственные лица) - собственные векторы ковариационной матрицы распределения изображений лиц в пространстве изображений. Каждое изображение лица может быть представлено как линейная комбинация eigenfaces.

Чтобы сократить размерность, можем оставлять только M **наиболее значимых** собственные векторы (тех, что соответствуют наибольшим собственным значениям).

Eigenfaces (алгоритм)

Инициализация:

- Собрать тренировочный (train) датасет из изображений лиц;
- Вычислить eigenfaces, оставив только M наиболее значимых. Эти векторы являются базисом нового **пространства лиц** (face space).
- Вычислить проекции всех изображений из тренировочного датасета на пространство лиц (их координаты в этом пространстве).

Распознавание:

- Вычислить проекцию нового изображения на пространство лиц;
- Определить, является ли это изображение лицом (проверить, достаточно ли близко оно к пространству лиц);
- Если это лицо, найти ближайшее изображение из тренировочного датасета в пространстве лиц (например, при помощи евклидова расстояния).

Eigenfaces (examples)



Eigenfaces (examples)



Eigenfaces (технически)

- Каждое изображение размера $N \times N$ пикселей представляется в виде точки (вектора) в N^2 -мерном пространстве. Каждое измерение - яркость пикселя (в ч/б).
- Пусть у нас есть фотографии (векторы) $\Gamma_1, \Gamma_2, \dots, \Gamma_M$. "Средняя" фотография - $\Psi = \frac{1}{M} \sum_{n=1}^m \Gamma_n$. Отнормируем: $\Phi_i = \Gamma_i - \Psi$. Векторы Φ_i составляют матрицу $A = [\Phi_1, \Phi_2, \dots, \Phi_M]$, ковариационную матрицу для которой мы и будем находить.
- Получаем собственные векторы - **eigenfaces** - u_1, \dots, u'_M , где M' - требуемая размерность нового пространства (гиперпараметр).

Eigenfaces (технически)

Пусть Γ - новая фотография. Тогда находим координаты в полученном пространстве по формуле:

$$\omega_k = u_k^T (\Gamma - \Psi).$$

Получаем координаты фотографии в face space:

$$\Omega^T = [\omega_1, \omega_2, \dots, \omega_M']$$

Находим из ранее известных классов такой класс k , для которого минимально:

$$\varepsilon_k = \|\Omega - \Omega_k\|^2, \text{ где } \Omega_k \text{ - координаты } k\text{-ого класса.}$$

Также для предварительной проверки близости к пространству лиц можно рассчитать:

$$\varepsilon^2 = \|\Phi - \Phi_f\|^2, \text{ где } \Phi = \Gamma - \Psi, \Phi_f \text{ - её проекция на face space.}$$

Эксперименты

Был собран датасет из фотографий 16 белых мужчин. Для каждого сделано 27 фотографий - 3 степени освещения, 3 степени наклона головы и 3 масштаба. Также эти фотографии были искусственно растянуты и сжаты в 6 расширений.

Deformation	Accuracy
Освещение	96%
Наклон	85%
Масштаб	64%

Таблица: Эксперимент без верхней границы отклонения

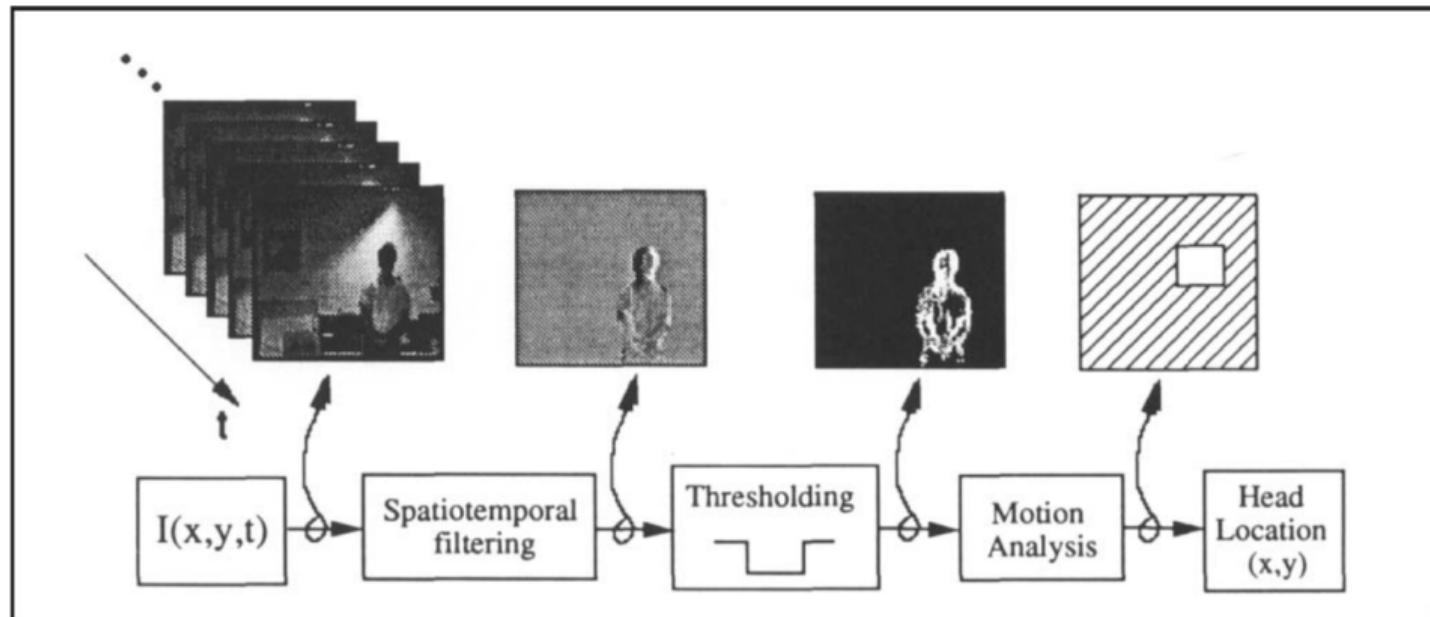
Эксперименты

При выставлении верхнего порога отклонения ε_k до такого уровня, который обеспечивал бы 100% accuracy, нераспознанными оказались:

Deformation	Unknown
Освещение	19%
Наклон	39%
Масштаб	60%

Таблица: Эксперимент с верхней границей отклонения

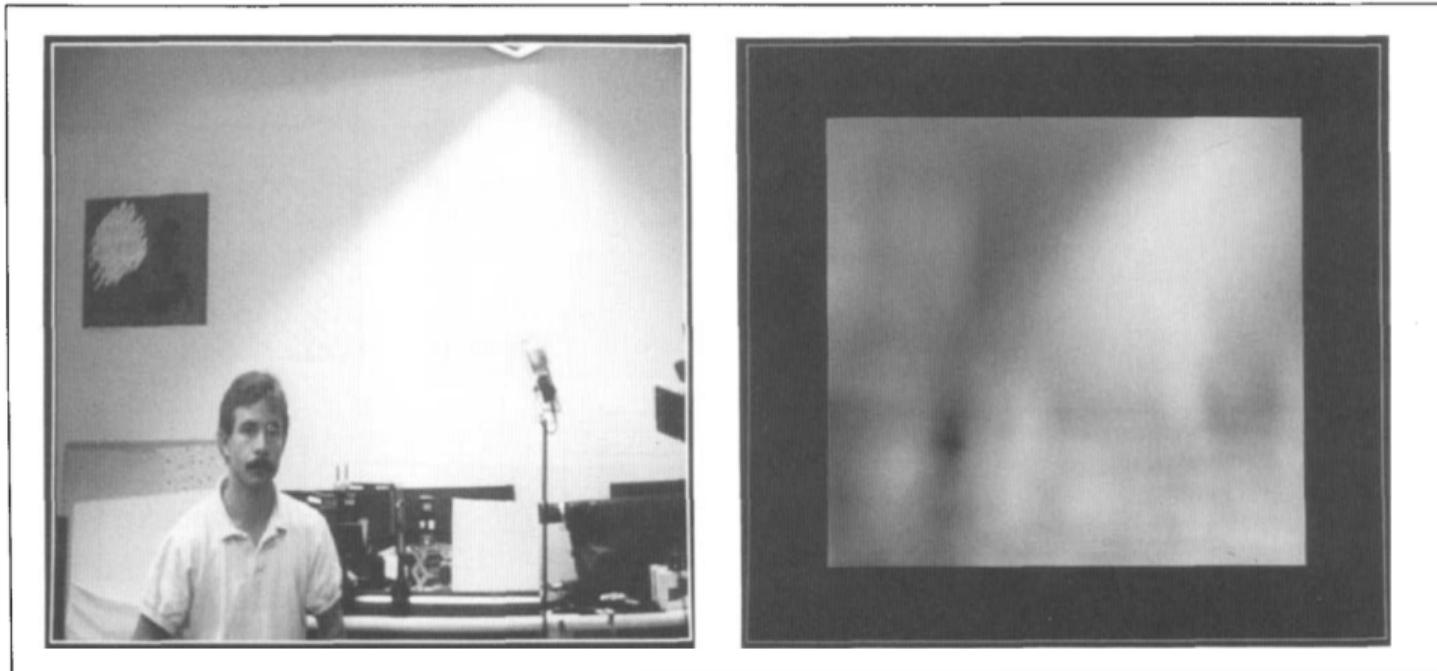
Face detection: Способ 1. Motion detection



Face detection: Способ 1. Motion detection



Face detection: Способ 2. Face space



Преимущества

- Переход от “feature-based” подхода (поиск глаз, носа, рта) к “appearance-based” (анализ целого изображения).
- Идея представить лицо как точку в пространстве главных компонент (face space) — принципиально новый взгляд, ставший основой десятков последующих работ.
- В отличие от сложных нейросетевых и геометрических моделей того времени, метод eigenfaces: использует простую линейную алгебру (нахождение собственных векторов ковариационной матрицы); позволяет выполнять обучение и распознавание в реальном времени (по меркам 1991 года); требует относительно небольшого объёма вычислений и памяти.
- Авторы опираются на данные нейрофизиологии (Perrett, Rolls и др.), связывая метод с возможными механизмами восприятия лиц у человека и приматов.

Недостатки

- Даже небольшое изменение угла поворота, яркости или расстояния вызывает заметное падение точности.
- Система не отделяет лицо от фона: ковариационная матрица строится по всему изображению. Из-за этого: фон влияет на вычисление собственных векторов; прически, очки или тени искажают результат.
- Метод РСА описывает данные линейной комбинацией базисных векторов. Но вариации лиц (мимика, выражение, тень) — нелинейные. Следовательно, РСА не может корректно сформировать сложные деформации.
- В оригинальных экспериментах использовалась база всего из 16 человек (2500 изображений). Современные стандарты требуют гораздо большего разнообразия. Недостаток: результаты проверены на малом и искусственно контролируемом наборе данных; отсутствует статистическая проверка обобщающей способности модели.

В настоящее время

Несмотря на то что метод Eigenfaces, предложенный Тёрком и Пентлендом в 1991 году, уже давно уступил место более совершенным подходам глубокого обучения, он по-прежнему сохраняет важное значение в теоретическом и образовательном контексте. Этот алгоритм положил начало целому направлению — статистическому анализу изображений лиц и построению признаков на основе подпространств, — что впоследствии стало концептуальной основой для таких моделей, как Fisherfaces, Laplacianfaces, Local Binary Patterns, а позднее — для deep face embeddings (FaceNet, DeepFace, ArcFace).