

1. Типы представлений изображений: низкоуровневые, контурные, структурные, иерархические

1. Представления в виде необработанных данных

Представление в виде необработанных данных — набора результатов физических измерений. В качестве физической величины чаще всего выступает интенсивность электромагнитного излучения.

Расположение пикселей и размерность их массива зависит от организации среды (например трехмерные изображения могут быть представлены массивом вокселей).

Пиксельные представления содержат все имеющуюся информацию, однако не удобны для автоматического анализа и не обеспечивают экономного хранения данных. Такие представления редко используются для распознавания или сопоставления.

2. Низкоуровневые представления

Представления изображений в качестве элементов математических пространств. Позволяет распространить на них все инструменты работы с такими объектами.

а) **Функциональные модели** — представления изображений как дискретных функций, например в Гильбертовом пространстве

$$f: G \rightarrow V, G \subseteq R^n, V \subseteq R^m$$

где G — область определения функции, V — область ее значений, n — размерность изображения, m — размерность вектора физических величин для каждой точки.

При таком представлении пространственные преобразования изображений (масштабирование, вращение, перенос) осуществляется при помощи смены системы координат.

$$f_2(x) = f(g(x)), g: G \rightarrow G_2$$

Преобразования яркости может быть представлено в виде некоторой функции от изображения.

$$f_2(x) = h(f(x)), h: V \rightarrow V_2$$

Фильтрация изображений в таком случае может быть представлена как операция линейной свертки изображения с некоторым ядром.

$$f_2 = T(f) = f * \varphi$$

Возможна также смена базиса (например, при помощи преобразования Фурье)

$$F(\omega) = \int_{-\infty}^{\infty} f(x) e^{-i\omega x} dx$$

Такие представления позволяют ввести понятие инварианта (например при необходимости работы с изображениями, инвариантными к конкретным типам преобразований).

б) **Вероятностные модели** — представление изображений как случайных функций. Берут свое начало из теории цифровой обработки сигналов.

Изначально использовались для задач подавления шума. Сейчас широко применяются для описания текстуры, реставрации изображений и т. д. Главный вопросом является — существует ли общая статистическая модель изображения? Ответ ищется в рамках порождающих статистических моделей изображений, в которых предполагается наличие набора скрытых переменных $x = (x_1, \dots, x_N)$ с заданной плотностью распределения априорных вероятностей $p(f|x)$, где f — изображение. Скрытые переменные должны каким-то образом описывать содержание сцены. В рамках этого подхода к представлению изображений применяются всевозможные методы статистического анализа сцен: построение гистограмм яркости и т. д.

3. Контурные

Под контуром обычно понимается местоположение локального изменения или резкого

перепада яркости на изображении. Они обычно соответствуют границам объектов. Построение контуров можно осуществлять либо глобально, с использованием разбиения изображения на однородные области, либо локально, используя анализ градиентного поля.

- а) контур является концентратором информации об изображении;
- б) контур полностью характеризует форму объектов на изображении;
- в) контуры устойчивы на изображениях от разных сенсоров и полученные в разное время;
- г) контурные точки — незначительная часть всех точек, что позволяет экономить ресурсы.

4. Структурные

а) **Признаковые представления**, как результат отклика некоторых шаблонов в точке. Это могут быть, например, перепады яркости, углы, отрезки определенного вида, пятна, полосы и т. д. Также используются детекторы признаков, предназначенные для узких областей (например распознавание лиц). (*Дискриминантная теория распознавания образов, использующая векторы признаков*)

б) **Структурные элементы**, строящиеся на основе контурных представлений. Их построение решается при помощи, например, сегментации контуров. В качестве таких элементов могут выделять отрезки прямых линий, дуги окружностей, эллипсы и т. д. (*Синтаксическая теория распознавания образов, использующие формальные грамматики, алфавиты символов и т. д.*)

в) **Составные структурные элементы**

5. Иерархические

а) **По пространственному масштабу**. Иерархичность по масштабу постулируется как одно из фундаментальных свойств видимого мира. Такой подход заметно увеличивает скорость работы методов распознавания, однако возрастает вероятность получения неоптимального решения, т. к. используется не вся доступная информация.

б) **По уровням абстрактности привлекаемых представлений**. Это методы, привлекающие, как правило, несколько представлений (пиксельные, контурные, структурные и т. д.). Задача распознавания решается на низком уровне, затем с привлечением более высокоуровневых представлений и с уточнением на низкоуровневых.

Используется теория адаптивного резонанса Гроссберга, говорящая о том, что получив грубое решение на высоком уровне, мы потом можем уточнить его на более низких уровнях. Получила развитие в теории нейронных сетей.