

Простые методы кодирования изображений с использованием преобразований реализуют отображение входного изображения в пространство коэффициентов того или иного линейного ортогонального преобразования с последующим отбрасыванием или грубым квантованием малых по величине указанных коэффициентов в процессе передачи изображения по каналу связи.

В результате прямого двумерного линейного ортогонального преобразования цифрового изображения $f(j, k)$ размером $N \times N$ образуется матрица преобразованного изображения того же размера. Элементы этой матрицы определяются следующим образом:

$$\hat{f}(m, n) = \sum_{j=1}^N \sum_{k=1}^N f(j, k) a(j, k; m, n), \quad (1)$$

где $a(j, k; m, n)$ – ядро прямого преобразования. Исходное изображение получается путем выполнения обратного преобразования вида:

$$f(j, k) = \sum_{m=1}^N \sum_{n=1}^N \hat{f}(m, n) b(j, k; m, n), \quad (2)$$

где $b(j, k; m, n)$ – ядро обратного преобразования.

Если массивы f и \hat{f} развернуть по строкам (или столбцам) и представить в виде векторов \mathbf{f} и $\hat{\mathbf{f}}$ соответственно, то выражения (1) и (2) можно записать в следующем виде:

$$\hat{\mathbf{f}} = \mathbf{A} \mathbf{f}, \quad (3)$$

$$\mathbf{f} = \mathbf{B} \hat{\mathbf{f}}, \quad (4)$$

где \mathbf{A} и \mathbf{B} – матрицы прямого и обратного преобразований размером $N^2 \times N^2$ элементов. Заметим, что если элементы матрицы \mathbf{A} являются действительными числами, то справедливо равенство:

$$\mathbf{B} = \mathbf{A}^T,$$

где «T» – знак транспонирования матрицы.

Нетрудно видеть, что выражения (3) и (4) могут быть реализованы однослойной нейронной сетью с линейной функцией активации нейронов, имеющей N^2 входов и N^2 выходов.

Известно, что наиболее оптимальным преобразованием цифрового изображения при его кодировании является преобразование Карунена-Лоэва. Однако выразить в явном виде ядро этого преобразования не всегда удается. Поэтому наиболее часто для целей кодирования цифровых изображений используется косинусное преобразование. Это связано с тем, что данное преобразование является весьма хорошим приближением к преобразованию Карунена-Лоэва. Вместе с тем, следует отметить, что преобразование Карунена-Лоэва может быть получено с использованием однослойной обучающейся нейронной сети с линейной функцией активации нейронов.

С целью ускорения реализации на обычных последовательных вычислительных средствах линейных ортогональных преобразований разработаны специальные алгоритмы быстрых преобразований. В нейросетевом представлении быстрые алгоритмы линейных ортогональных преобразований представляют собой многослойные нейронные сети с линейной функцией активации. На рис. 1 приведен пример нейросетевого представления простого (а) и быстрого (б) одномерного преобразования Адамара для последовательности из четырех отсчетов $f(j)$.

Из рисунка следует, что для реализации простого преобразования Адамара требуется выполнить 12 операций сложения и вычитания, а быстрого преобразования – 8 операций.

В научных публикациях даже появился специальный термин “быстрые нейронные сети”. Однако, с точки зрения нейрокомпьютинга такое представление вычислительного алгоритма не всегда является целесообразным, поскольку многослойные нейронные сети с линейной функцией активации легко преобразуются к однослойным нейронным сетям. Потенциально это позволяет существенно повысить оперативность необходимых вычислений на нейрокомпьютерах. Данное утверждение базируется на том, что нейроны одного слоя по определению нейронной сети функционируют параллельно. Поэтому в пределе оперативность реализации рассматриваемых преобразований на нейровычислительных средствах определяется оперативностью выполнения отдельным процессором N^2 операций сложения и умножения. При использовании однопроцессорных универсальных вычислительных средств для реализации линейных ортогональных преобразований по формуле (1) потребуется примерно N^4 операций. При реализации быстрых алгоритмов на тех же вычислительных средствах потребуется [1] примерно $2N^2 \log_2 N$ операций. Таким образом, нейрокомпьютерная реализация линейных ортогональных преобразований потенциально

обеспечивает максимально возможную оперативность их реализации. Кроме того, нейрокompьютеры обеспечивают принципиальную реализуемость на них параллельной блочной обработки изображений, т.е. все блоки изображения, для каждого из которых вычисляются преобразования (1) или (2), могут обрабатываться параллельно. Это в значительной степени повышает оперативность обработки изображений на нейрокompьютерах.

В реальных нейрокompьютерах одним процессором могут обрабатываться вычисления, реализуемые группой нейронов. В этих случаях реализация быстрых линейных ортогональных преобразований на нейрокompьютерах также может оказаться полезной.

Существует множество различных схмотехнических решений в области нейрокompьютеров, которые наряду с характеристиками используемой элементной базой и определяют производительность конкретных нейрокompьютеров.

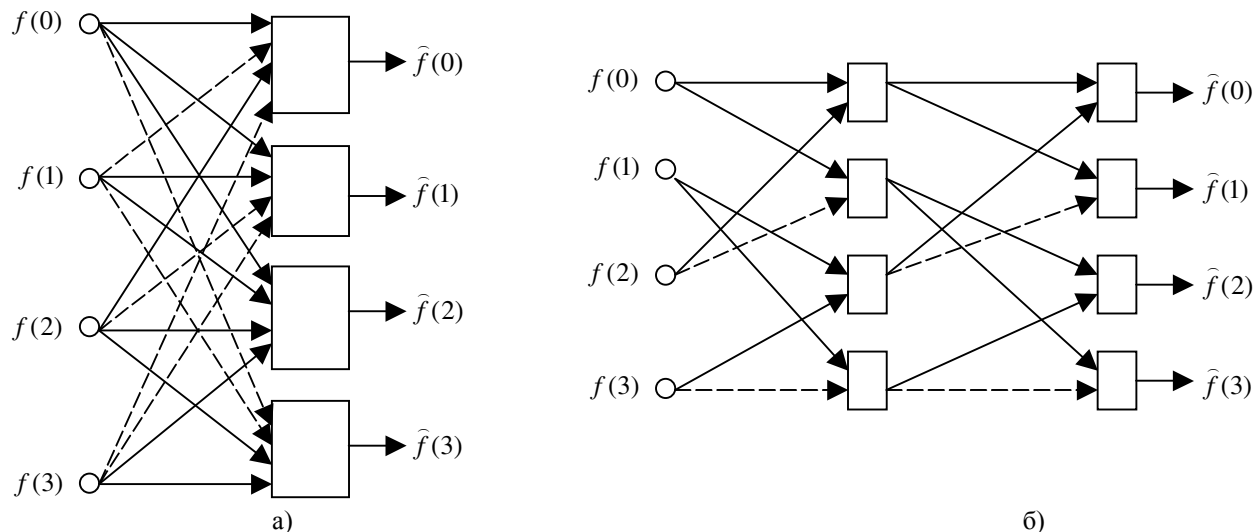


Рис. 1. Нейросетевой представление одномерного преобразования Адамара при $N=4$: а) однослойная нейронная сеть с линейной функцией активации нейронов, реализующая простое преобразование Адамара; б) двухслойная нейронная сеть с линейной функцией активации нейронов, реализующая быстрое преобразование Адамара. Пунктирные линии обозначают связи с весовыми коэффициентами, равными -1 , а сплошные – связи с весовыми коэффициентами, равными 1 .

Любое линейное ортогональное преобразование фиксировано, поэтому никакие адаптивные процедуры к нему не применимы. Однако адаптивным может быть квантование спектральных коэффициентов преобразования в процессе передачи преобразованного изображения по каналу связи. Поскольку преобразование изображения осуществляется, как правило, поблочно, то квантование также осуществляется поблочно посредством перехода с одной матрицы распределения битов на другую.

Эти матрицы будучи однажды определенными используются для всех изображений и хранятся как на передающем, так и приемном устройствах.

В пределах одного класса блоки квантуются с использованием одной и той же матрицы распределения битов.

Классификация блоков преобразованных изображений может осуществляться с использованием самообучающихся нейронных сетей. При этом в качестве признаков могут выступать как сами коэффициенты косинусного преобразования, так и признаки, определенные на множестве указанных коэффициентов, например, различные текстурные меры. Базовой основой для построения процедур самообучения может служить минимизация функции среднего риска.

Более общим нейросетевым подходом к реализации сжатия изображений и их восстановления может служить подход, основанный на использовании многослойных обучающихся нейронных сетей, реализующих нелинейное преобразование изображений при их сжатии и восстановлении.

Литература

1. Прэт У. Цифровая обработка изображений. В 2-х кн. – М.: Мир, 1982.