

Топологический подход к распознаванию линейных фигур

М. Марьянович, Р. Томович, С. Станкович

С комбинаторной точки зрения форма является конечной совокупностью дуг на плоскости в некоторой системе координат, пересекающихся только в концевых точках. С топологической точки зрения, форма – это пространство, лежащее в основании этой совокупности дуг.

Следуя гештальт-психологии, основную структуру формы рассматривают с точки зрения пространственного расположения дуг и их ориентации. Таким образом, две формы считаются эквивалентными, если они гомеоморфно отображаются друг на друга с помощью некоторого гомеоморфизма, сохраняющего их структуру.

Гештальт (тип формы) определяется как стандартный эталон формы и в равной степени может быть представлен любой из эквивалентных форм, рассматриваемых как эталон, искаженный глобальным шумом: отсутствием регулярности в форме дуг и дисбалансом в структуре формы. Например, цифра «2» определяется как любая форма, эквивалентная ее (печатному) эталону.

С формальной точки зрения, дополнительные свойства, налагаемые на дуги, позволяют взглянуть на форму как на график многозначной функции, и использовать непрерывность этой функции. Далее, форме сопоставляется последовательность чисел (дискриминант непрерывности), и мы доказываем их связь с эйлеровой характеристикой формы. Матрица, сопоставленная форме, выявляет ее глобальную структуру. Применение этих последовательностей и матриц ведет к арифметизации проблемы классификации. Мы надеемся эффективно использовать это в практических процедурах распознавания.

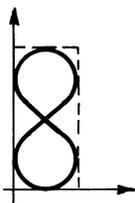
1. Введение

Когда мы воспринимаем некоторый объект, одна его часть привлекает наше внимание, в то время как все остальное остается на заднем плане. Например, при восприятии следующей формы числа «8» нарисованная линия



выделяется по отношению к заднему плану, которым является окружающая линия части страницы (или страница целиком).

Принимая во внимание обычное направление чтения и письма (в западной культуре), прямоугольная рамка (показанная на следующем рисунке) является минимальным задним планом формы, а плоскость, ориентированная посредством заданной на ней координатной системы, – максимальным задним планом.



Рамка выступает здесь как разделитель, так что форма воспринимается нами как организованное целое, в то время как координатная система заставляет нас смотреть на форму как на график некоторой многозначной функции.

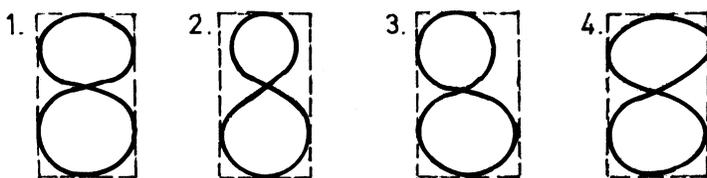
Все мы учились писать буквы и цифры по одинаковым стандартным эталонам. Тем не менее, наш почерк разнится очень сильно, отражая индивидуальный стиль письма. Мы легко читаем рукописный текст, написанный достаточно аккуратно. Мы оцениваем такой почерк, говоря, что он «хороший» или «плохой». Это означает, что стиль письма оказывает на нас эстетическое воздействие. Если текст читается легко, мы рассматриваем такой почерк как «хороший». В этом случае мы обнаруживаем, что все буквы различаются между собой

и сохраняют структуру своих эталонов очень хорошо.

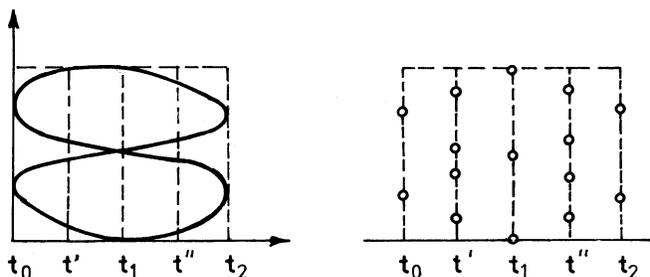
Предположим теперь, что форма задана посредством стандартного эталона. Какая из ее реализаций является «хорошей», а какая – «плохой»? Если вы знакомы с трудами М. Вертхаймера [9] или В. Кёлера [1], вы без труда склонитесь к тому, чтобы следовать идеям Гештальт-психологии.

Хотя это и не является достаточно объективным критерием, тем не менее «хорошая» форма предполагает наличие таких свойств, как симметричность, сбалансированность и полноту. Таая форма легко идентифицируется по своему эталону. «Плохая» форма искажена шумом (присутствием неструктурных свойств), причиной которого является «плохая» реализация формы (отсутствие регулярности, выпадение отдельных частей и т.п.). Идентификация «плохой» формы с помощью ее эталона обычно является более сложной процедурой (некоторые очень оригинальные стилизации букв, которые можно встретить, например, в рекламе, можно также отнести к «плохим» формам).

Вернемся снова к примеру с цифрой «8» и рассмотрим несколько ее реализаций.



Рассмотрим увеличенную копию реализации 1 и введем для нее координатную систему. Посмотрим на эту форму как на график определенной таким образом многозначной функции F .



Значения функции F таковы: в точке t – множество из двух элементов, в точке t' – множество из четырех элементов и т.д. Когда мы движемся от t' к t_0 , четырехэлементное множество $F(t)$ становится близким к двухэлементному множеству $F(t_0)$. В этом случае мы говорим, что F непрерывна справа в точке t_0 . Исследуя далее функцию F , мы с легкостью видим, что в этом смысле она непрерывна во всех точка отрезка $[t_0, t_1]$.

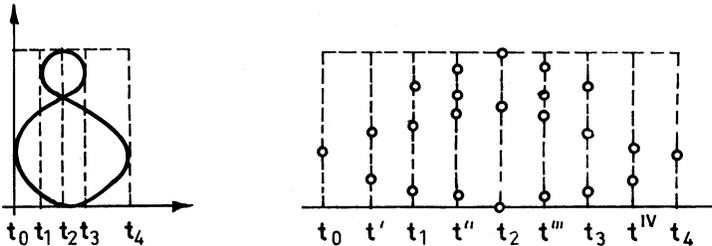
Заметим, что отрезок $[t_0, t_1]$ является областью определения функции F , и что он разделяется точкой t' на два подотрезка. Можно видеть, что мощность множества $F(t)$ одинакова для всех внутренних точек этих подотрезков. Обозначим через (n_t) мощность множества $F(t)$, когда t является внутренней точкой подотрезка, в противном случае обозначим мощность $F(t)$ через n_t . Теперь сопоставим рассматриваемой форме последовательность

$$\tau_1 = n_{t_0}(n_{t'})n_{t_1}(n_{t''})n_{t_2}$$

В нашем конкретном случае мы имеем

$$\tau_1 = 2(4)3(4)2.$$

То же самое сделаем для реализации 2. Таким образом, мы имеем



Исследуя непрерывность F , мы находим, что при стремлении t' к t_1 слева $F(t')$ не приближается к $F(t_1)$. Таким образом, F не является непрерывной слева в точке t_1 . Соответствующая последовательность для этого случая выгладит так:

$$\tau_2 = 1(2) * 3(4)3(4)3 * (2)1,$$

где звездочкой обозначаются места разрывов функции F .

Аналогично, в случае 3 мы имеем

$$\tau_3 = 2(4)3(4)3 * (2)1,$$

а в случае 4 –

$$\tau_4 = 2(4)3(4)3.$$

Просуммировав числа в этих последовательностях с чередующимися знаками, мы получим:

1. $2 - 4 + 3 - 4 + 2 = -1$
2. $1 - 2 + 3 - 4 + 3 - 4 + 3 - 2 + 1 = -1$
3. $2 - 4 + 3 - 4 + 3 - 2 + 1 = -1$

Каждая форма может рассматриваться как топологическое пространство (пространство, лежащее в основании формы) и число -1 является Эйлеровой характеристикой этих пространств.

В случае 4, мы имеем

$$4. \quad 2 - 4 + 3 - 4 + 3 = 0,$$

и Эйлерова характеристика этого пространства равна 0. Следовательно, это пространство не является топологически эквивалентным пространству «цифра восемь».

Является ли четвертая реализация очень «плохой»? В этом случае лучше сказать, что мы имеем реализацию формы, отличную от той, что нам дана эталоном 1.

В первых трех случаях чем проще соответствующая случаю последовательность, тем «лучше» форма. Например, форма для случая 2 «хуже» формы для случая 3 (она менее сбалансирована).

В случае 4 мы склоняемся к тому, что перед нами цифра «восемь». Но в этом случае мы рассматриваем ее как стимул. А наша позиция в этом случае – рассматривать форму как некий отклик. Более точно, форма представляет что-то, что уже воспринято нами. И мы представляем эти наши восприятия используя коды непрерывности на плоскости.

Чтобы идеи, коротко изложенные в этом введении, сделать более строгими, нам необходимо ввести некоторые основные топологические понятия.

2. Хаусдорфова метрика

Вспомним, что *метрическим пространством* называется пара (M, d) , где M – непустое множество, а $d : M \rightarrow [0; +\infty)$ – функция, такая что для любой тройки точек $x, y, z \in M$ выполнено:

- 1) $d(x, y) = 0 \Leftrightarrow x = y$,
- 2) $d(x, y) = d(y, x)$,
- 3) $d(x, y) + d(y, z) \geq d(x, z)$.

Далее, последовательность (x_n) точек из M сходится к точке $x_0 \in M$, если:

$$(\forall \varepsilon > 0)(\exists m)(\forall n)n \geq m \Rightarrow d(x_n, x_0) < \varepsilon.$$

Если (M_1, d_1) и (M_2, d_2) – метрические пространства, то отображение $f : M_1 \rightarrow M_2$ *непрерывно* в точке $t_0 \in M_1$, если для любой последовательности (x_n) точек из M_1 выполнено:

$$x_n \rightarrow t_0 \Rightarrow f(x_n) \rightarrow f(t_0).$$

Говорим, что отображение f *непрерывно*, если оно непрерывно в любой точке $x \in M_1$.

Открытым шаром $B(x_0, r)$ в пространстве (M, d) с центром в точке x_0 и радиусом $r > 0$ назовем следующее подмножество M :

$$B(x_0, r) = \{x \in M \mid d(x_0, x) < r\}.$$

Подмножество $U \subseteq M$ будем называть *открытым*, если для любой точки $x \in U$ существует $r > 0$, такое что $B(x, r) \subseteq U$ (или, что эквивалентно, если U является объединением открытых шаров). Дополнение к U – множество $F = M \setminus U$ – называется *замкнутым* (или, соответственно, множество F называется *замкнутым*, если для каждой последовательности точек (x_n) из F и точки $x_0 \in M : x_n \rightarrow x_0 \Rightarrow x_0 \in F$). Подмножество $K \subseteq M$ называется *компактным*, если каждая последовательность точек (x_n) из F имеющая сходящуюся в F подпоследовательность $(x_{k(n)})$ (то есть пределом этой последовательности является точка из F). Вспомним, что

компактные множества в Евклидовых пространствах – это в точности те множества, которые являются замкнутыми и ограниченными.

Расстояние между двумя непустыми подмножествами A, B из M определяется как

$$\text{dist}(A, B) = \inf\{d(x, y) \mid x \in A, y \in B\}.$$

Когда A и B компактны, инфимум достигается в некоторой паре точек (x_0, y_0) , и тогда $\text{dist}(A, B) = d(x_0, y_0)$.

Функция dist не удовлетворяет свойствам 1. и 3. метрики, даже в случае, когда A и B являются конечными.

Предположим, что A и B – непустые компактные подмножества множества M . Для $x \in A, y \in B$ рассмотрим величины

$$\rho(A, B) = \sup_x \inf_y d(x, y), \quad \rho(B, A) = \sup_y \inf_x d(x, y).$$

Эти два числа могут часто быть разными, например в случае, когда B является собственным подмножеством A , имеем $\rho(A, B) > 0$ и $\rho(B, A) = 0$. (Возьмите $A = \{0, 1\}, B = \{1\}$ и проверьте, что $\rho(A, B) = 1, \rho(B, A) = 0$).

Большее из двух чисел

$$d(A, B) = \max\{\rho(A, B), \rho(B, A)\}$$

является метрикой на множестве всех непустых компактных подмножеств множества M , которое обозначим через $\text{exp}(M)$. Эта метрика называется *Хаусдорфовой метрикой*, а пространство $(\text{exp}(M), d)$ называется гиперпространством компактных подмножеств M . Рассматривая каждую точку $x \in M$ как подмножество $\{x\} \subseteq M$ и обозначив $d(x, y) = d(\{x\}, \{y\})$, мы увидим, что пространство M может рассматриваться как «маленькое» подпространство пространства $\text{exp}(M)$ – «большого» расширения M .

Следующие полезные утверждения легко доказываются:

Утверждение 1.

$$d(A, B) < \varepsilon \Leftrightarrow \begin{cases} (I) & (\forall x \in A)(\exists y \in B)d(x, y) < \varepsilon \\ (II) & (\forall y \in B)(\exists x \in A)d(x, y) < \varepsilon \end{cases}$$

Иначе говоря, $d(A, B) < \varepsilon$ тогда и только тогда, когда для любой точки $x \in A$ существует точка $y \in B$, такая что $d(x, y) < \varepsilon$ и для любой точки $y \in B$ существует $x \in A$, такая что $d(x, y) < \varepsilon$.

Пусть A – непустое компактное подмножество множества M и $\varepsilon > 0$. Конечное подмножество $F \subset A$ называется ε -сетью множества A , если для любой точки $x \in A$ существует точка $y \in F$, такая что $d(x, y) < \varepsilon$. Хорошо известно, что для любого множества A и любого $\varepsilon > 0$ существует ε -сеть A .

Поскольку $\rho(A, F) < \varepsilon$ и $\rho(F, A) = 0$, мы имеем, что $d(A, F) < \varepsilon$. Это означает, что каждое $A \in \text{exp}(M)$ может быть сколь угодно хорошо аппроксимировано конечным подмножеством. В этом смысле, конечные подмножества M играют роль рациональных чисел во множестве действительных чисел. Запишем это свойство конечных множеств на формальном языке.

Утверждение 2. *Множество всех конечных подмножеств M всюду плотно в $\text{exp}(M)$.*

Заметим, что A является подмножеством M и элементом множества $\text{exp}(M)$. Многозначное отображение $F : M_1 \rightarrow M_2$ ставит в соответствие каждой точке $x \in M_1$ подмножество $F(x)$ множества M_2 . Будем предполагать, что $F(x)$ компактно для любой точки $x \in M_1$. Тогда соответствие $x \rightarrow F(x)$ может рассматриваться как однозначное отображение M_1 в $\text{exp}(M_2)$. К отображению $F : M_1 \rightarrow \text{exp}(M_2)$ мы применяем определение непрерывности, данное нами для однозначного случая. В этом смысле мы говорим о *Непрерывности многозначных отображений*.

Для интерпретации можно рассмотреть примеры, данные во введении. Нестрогие выражения «близко» и «далеко», которые мы использовали там, приобретают теперь строгое логическое значение. Заметим, что для $m < n$ и фиксированного n -точечного множества F_0 , очевидно, что ни одно m -точечное множество F не является близким к F_0 . Заметим также, что пунктирная линия является естественным Хаусдорфовым приближением непрерывной линии. Когда точки

пунктира случайно распределены, их увеличение улучшает аппроксимацию. Но эта свойство является чисто метрическим! Обратное, более грубая аппроксимация может показывать лучше глобальные свойства формы.

3. Некоторые основные топологические понятия

Вспомним, что два метрических пространства M_1 и M_2 *топологически эквивалентны*, или *гомеоморфны*, если существует отображение $f : M_1 \rightarrow M_2$, которое (I) взаимнооднозначно и является отображением на и (II) оба отображения f и f^{-1} непрерывны. Такое отображение называется гомеоморфизмом, а про два гомеоморфных пространства говорят, что они имеют *один и тот же топологический тип*. Далее будем писать $M_1 \approx M_2$. Свойство, которым обладают все пространства одного топологического типа, называется *топологическим свойством* или *топологическим инвариантом*. Например, компактность является топологическим свойством.

Непрерывное отображение $u : [0, 1] \rightarrow M$ называется *путем* из точки $x = u(0)$ в точку $y = u(1)$. Подмножество $A \subseteq M$ называется *линейно связным*, если для любых двух точек $x, y \in A$ существует путь $u : [0, 1] \rightarrow A$, такой что $u(0) = x$ и $u(1) = y$. *Компонентой связности* называется наибольшее линейно связное подмножество M . Линейная связность является также топологическим свойством.

Два непрерывных отображения $f_0, f_1 : M_1 \rightarrow M_2$ называются *гомотопными*, если существует непрерывное отображение $F : M_1 \times [0, 1] \rightarrow M_2$, такое что $F(x, 0) = f_0(x)$ и $F(x, 1) = f_1(x)$. Далее будем писать $f_0 \cong f_1$. Два пространства M_1 и M_2 *гомотопно эквивалентны* (*гомотопны*), если существуют два непрерывных отображения

$$M_1 \begin{array}{c} \xrightarrow{f} \\ \xleftarrow{g} \end{array} M_2$$

такое что $f \circ g \cong 1_{M_1}$ и $g \circ f \cong 1_{M_2}$. Далее будем писать $M_1 \cong M_2$ и говорить, что M_1 и M_2 имеют *один гомотопный тип*. Свойство,

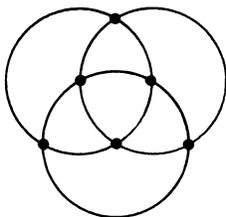
которым обладают все пространства одного гомотопного типа, называется *гомотопным свойством* или *гомотопным инвариантом*.

Два топологически эквивалентных пространства являются гомотопно эквивалентными. Следовательно, гомотопный инвариант является также и топологическим инвариантом. Например, линейная связность является гомотопным инвариантом, а компактность — нет.

Эти понятия в общем случае имеют очень сложное значение, но распространим наши рассуждения на маленький класс регулярных объектов на плоскости. Внутри этого класса значения этих понятий становятся проще и более описательными.

Пусть E^2 — Евклидова плоскость, а $f : [0, 1] \rightarrow E^2$ — взаимно-однозначное отображение. В записи $f(t) = ((f_1(t), f_2(t)))$ будем подразумевать, что f_1, f_2 являются также непрерывно дифференцируемыми и что для каждого t $(f_1'(t), f_2'(t))^2 \neq 0$. Далее, множество $\{f(t) | t \in [0, 1]\}$ называется *гладкой дугой* с концами $f(0)$ и $f(1)$. Множество $\{f(t) | t \in (0, 1)\}$ называется *внутренностью* этой дуги.

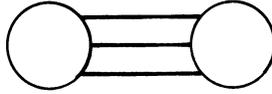
Связный 1-комплекс во множестве E^2 — это подмножество E^2 , которое может рассматриваться как конечное объединение гладких дуг, имеющих непересекающиеся внутренности. Например, следующие три окружности образуют 1-комплекс, рассматриваемый как объединение двенадцати дуг, имеющих в качестве концевых вершин точки пересечения окружностей.



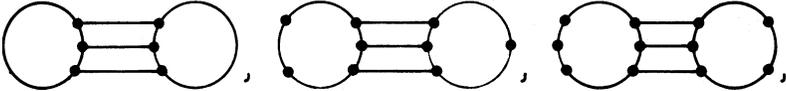
Точка в E^2 называется *связным 0-комплексом* (здесь k в сложном k -комплексе обозначает его размерность).

Связный 1-комплекс — это подмножество E^2 и, таким образом, подпространство E^2 , которое называется *пространством*, лежащим в основании комплекса.

То же самое подмножество может рассматриваться как 1-комплекс разными способами. Например, подмножество



может рассматриваться как:



В каждом случае, если через e обозначить число точек, которые являются концевыми, а через a — число дуг, то

$$e - a = 6 - 9 = 9 - 12 = 11 - 14 = -3$$

Числа e , a зависят от способа представления подмножества в качестве 1-комплекса, а разность $e - a$ не зависит от этого. Число $e - a$ зависит только от типа гомотопии лежащего в основании комплекса пространства и называется *Эйлеровой характеристикой пространства*.

(Читатель, хорошо знакомый с топологией, вспомнит, что эйлерова характеристика определяется для пространств, являющихся конечными CW -комплексами. Для такого пространства X число

$$\chi(X) = \sum_{\nu=0}^n (-1)^\nu \beta_\nu,$$

где величины β_ν суть числа Бэтти, называется Эйлеровой характеристикой пространства X и это число является гомотопным инвариантом также как и вся формула в этом определении. Если α_ν — число ν -ячеек X , то α_ν не является топологической характеристикой пространства, но следующее равенство:

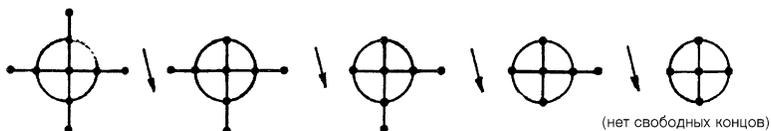
$$\sum_{\nu=0}^n (-1)^\nu \alpha_\nu = \chi(X)$$

имеет место.

Если X, Y являются конечными CW -комплексами и $Z = X \cap Y$ – подкомплекс X и Y , то

$$\chi(X \cup Y) = \chi(X) + \chi(Y) - \chi(Z).$$

1-комплекс является *свободным концом*, если он является концом в точности одной дуги комплекса. *Элементарным разрушением* назовем удаление из комплекса свободного конца и соответствующей ему дуги. Последовательность элементарных разрушений назовем *разрушением*. Разрушение интерпретируется рисунком



и все эти пространства являются гомотопно эквивалентными, потому что *разрушение сохраняет тип гомотопии* (и Эйлерову характеристику). После каждого элементарного разрушения, выражение $e - a$ меняется на $(e - 1) - (a - 1)$, но разность остается той же.

4. Дискриминант непрерывности формы

Пусть E^2 – Евклидова плоскость с фиксированной координатной системой (ось x пусть будет направлена вправо, а ось y – вверх).

Мы должны избежать рассмотрения таких затруднительных дуг, как

$$t \rightarrow (t^4 \sin 1/t, t).$$

Таким образом, гладкая дуга является *регулярной*, если когда бы она не пересекала вертикальную прямую $x = t$, пересечение является либо точкой, либо дугой целиком (в последнем случае дуга является замкнутым отрезком прямой $x = t$). Комплекс *регулярен*, если все его дуги регулярны.

Линейная форма (более короткая форма) в E^2 – это пространство, лежащее в основании регулярного связного 1-комплекса в E^2 .

Фрейм формы – это наименьший прямоугольник со сторонами, параллельными осям координат, содержащий форму

Связанная (с формой) координатная система – это параллельный перенос данной системы координат в левую нижнюю вершину фрейма.

Например,



– это две формы на плоскости этой страницы, если их рассматривать в качестве регулярных комплексов, они превращаются в



(жирные точки – концевые вершины).

Фрейм формы в связанной с ней координатной системе – это прямоугольник $[0, a] \times [0, b]$, где отрезок $[0, a]$ называется *областью определения* формы.

Пусть F – линейная форма, а $[0, a]$ – ее область определения. Для $t \in [0, a]$ пересечение формы F с прямой $x = t$ состоит из конечного числа компонент, являющихся либо точками, либо замкнутыми интервалами. Обозначим через $D(t)$ проекцию этого пересечения на ось y . Таким образом, соответствие $t \rightarrow F(t)$ является многозначной функцией $F : [0, a] \rightarrow E^1$ (где E^1 обозначает Евклидову прямую) и значение $F(t)$ является компактным множеством.

Не принимая во внимание, каким образом форма F рассматривается в качестве регулярного комплекса, она имеет конечное число концевых вершин и пересечение F с прямой $x = t$ содержит некоторые концевые вершины только для конечного числа моментов времени t . Обозначим их через

$$t_0 = 0 < t_1 \dots < t_k = a$$

Легко видеть, что для каждого отдельного интервала $[t_i, t_{i+1}]$ и каждого момента времени $t \in (t_i, t_{i+1})$, функция F непрерывна по t . Таким образом, единственно возможными точками разрывов функции F являются моменты времени t_i .

Обозначим через n_t число компонент $F(t)$. Заметим, что для каждого $t \in (t_i, t_{i+1})$, числа n_t равны. Для обозначения точек разрыв-

ности F мы добавим звездочки к n_{t_i} (сохраняя их уже зафиксированных ранее смысл). А именно, пусть

$$\tilde{n}_{t_i} = \begin{cases} n_{t_i}, & F \text{ непрерывна в точке } t_i \\ n_{t_i}, & F \text{ непрерывна справа и} \\ & \text{разрывна слева в точке } t_i \\ n_{t_i}^*, & F \text{ непрерывна слева и} \\ & \text{разрывна справа в точке } t_i \\ n_{t_i}^*, & F \text{ разрывна по обе стороны } t_i \end{cases}$$

Следуя уже описанному разбиению отрезка $[0, a]$

$$t_0 < t_1 < \dots < t_i < t_{i+1} < \dots < t_{k-1} < t_k,$$

которое мы связали с формой F , следующая последовательность

$$\tau_F = \tilde{n}_{t_0}(n_t^{(0)})\tilde{n}_{t_1} \dots \tilde{n}_{t_i}(n_t^{(1)})\tilde{n}_{t_{i+1}} \dots \tilde{n}_{t_{k-1}}(n_t^{(k-1)})\tilde{n}_{t_k},$$

котоую назовем *дискриминантом непрерывности* формы F .

Заметим, что последовательность τ_F имеет k членов в скобках и $(k+1)$ – без, то есть вместе – нечетное число $2k+1$ элементов. Эта последовательность связана с топологией формы F с помощью следующего утверждения.

Утверждение 3. Пусть $\chi(F)$ – Эйлерова характеристика линейной формы, а τ_F – ее дискриминант непрерывности. Тогда

$$n_{t_0} - n_t^{(0)} + n_{t_1} - \dots + n_{t_i} - n_t^{(i)} + n_{t_{i+1}} - \dots + n_{t_{k-1}} - n_t^{(k-1)} + n_{t_k} = \chi(F)$$

Доказательство. Будем использовать свойство

$$\chi(A \cup B) = \chi(A) + \chi(B) - \chi(A \cap B)$$

Эйлеровой характеристики. Согласно этому свойству, если $F = F_1 \cup \dots \cup F_m$ является объединением непересекающихся форм F_i , имеющих концевые вершины на прямых $x = t_i, i = 0, \dots, k$ и если

$$n_t(F) = n_t(F_1) + \dots + n_t(F_m),$$

то формула в утверждении выполнена для такой формы F тогда и только тогда, когда она выполнена для каждой формы $F_i, i = 1, \dots, m$.

Заметим, что пересечение $\tilde{F}(t) = \{(x, y) | x = t\} \cap F$ и $F(t)$ гомеоморфны (и гомеоморфизм является проекцией на ось y).

Теперь предположим, что F – такая форма, что $\tau_F = n_{t_0}$. Тогда либо F – либо точка, либо интервал, и $n_{t_0} = 1$. В то время как интервал может быть подвержен разрушению до одной из своих концевых вершин, мы имеем $\chi(F) = 1$. Таким образом, формула справедлива тогда, последовательность τ_F состоит из одного члена.

Чтобы доказать формулу в общем случае, используем индукцию по числу $2k + 1$ членов в τ_F (мы уже доказали формулу для $k = 0$).

Предположим, что формула верна для всех таких форм F , для которых τ_F имеет $2k - 1$ членов ($c k \geq 1$).

Пусть F – форма и пусть τ_F имеет ровно $2k + 1$ членов. Зафиксируем $t \in (t_{k-1}, t_k)$ и положим

$$F_0 = \{(x, y) \in F | x \leq t\}, F_1 = \{(x, y) \in F | t \leq x\}.$$

Тогда $F_0 \cap F_1 = \tilde{F}(t)$. Дуги в F_0 , имеющие концевую вершину в $F(t_{k-1})$ и другую (свободную) в $\tilde{F}(t)$ могут быть удалены, а F может быть разрушена до $\underline{F}_0 = \{(x, y) \in F | x \leq t_{k-1}\}$. Таким образом, $\underline{F}_0 \cong F_0$. Точно также $F_1 \cong \tilde{F}(t_k) = \underline{F}_1$. Поскольку \underline{F}_0 является объединением непересекающихся форм, дискриминанты непрерывности которых имеют максимум $2k - 1$ членов, после применения предположения индукции мы имеем

$$\chi(F_0) = \chi(\underline{F}_0) = n_{t_0} - n_t^{(0)} + n_{t_1} - \dots + n_{t_{k-1}}.$$

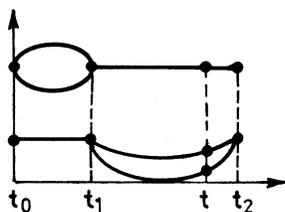
Поскольку

$$\chi(F_1) = \chi(\underline{F}_1) = n_{t_k}, \chi(\tilde{F}(t)) = n_t^{(k-1)},$$

мы находим

$$\chi(F) = \chi(F_0) + \chi(F_1) - \chi(\tilde{F}(t)) = (n_{t_0} - n_t^{(0)} + n_{t_1} - \dots + n_{t_{k-1}}) + n_{t_k} - n_t^{(k-1)}.$$

Это доказывает формулу в общем случае.



(Этот рисунок иллюстрирует доказательство)

5. Эталоны

Большую роль в мышлении играют символы, являющиеся в большей или меньшей степени стандартизованными формами. В качестве примера такой системы символов могут служить цифры

0, 1, 2, 3, 4, 5, 6, 7, 8, 9

написание которых можно встретить в простых учебниках по математике. Записав их дискриминант непрерывности, получим:

$$\tau_0 = 1(2)1, \tau_1 = 1(1) * 1, \tau_2 = 2(3)2, \tau_3 = 2(2) * 3(4)2$$

$$\tau_4 = 1(2) * 1(1)1, \tau_5 = 2 * (3)2, \tau_6 = 1(3)2, \tau_7 = 2(2)1$$

$$\tau_8 = 2(4)3(4)2, \tau_9 = 2(3)1$$

Полученные таким образом последовательности в трех случаях состоят из пяти членов. Можем ли мы сократить их, заменив фигуру на их хорошо распознаваемый вариант? Да, в случае первых двух таких цифр:

$$\begin{array}{cc} \mathbf{3} & \mathbf{4} \\ \tau_3 = \mathbf{3(4)2} & \tau_4 = \mathbf{1(2)_*1} \end{array}$$

и нет, в третьем случае (потому что пересечение двух линий должно быть конечной точкой). Заменим написания цифр 3 и 4, характерные для учебников по элементарной математике, на и предложенные варианты. Полученные таким образом 10 форм будут эталонами арабских цифр.

Определим *эталон* как форму, заданную в координатном представлении. Например, цифра 1 задается в координатах как

$$\{(x, y) | (9 \leq x < 10 \text{ и } y = x + 1) \text{ или } (x = 10 \text{ и } 0 \leq y \leq 2)\}$$

Несколько удовлетворительных уравнений подобного рода для дуг арабских чисел могут быть найдены. Таким образом, все эти эталоны могут быть заданы в координатах.

В качестве другого набора эталонов арабских чисел может служить следующее множество их реализаций в виде многоугольников:



Когда форма рассматривается как графический символ, ее размер не влияет на ее значение. Таким образом, мы не будем делать различия между формой F и ее аффинными образами, полученными в результате отображений вида $(x, y) \rightarrow (ax, ay)$ для $a > 0$ и $b > 0$.

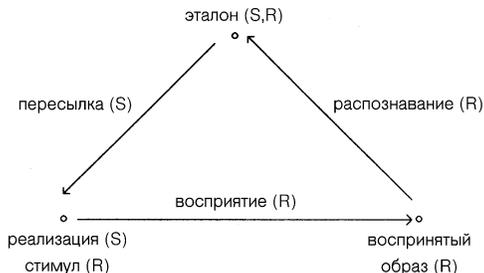
6. Общение между двумя людьми

В общении между людьми эталон может передаваться от человека к человеку с абсолютной точностью. На практике мы считаем, что эталоны передаются посредством их совершенных графических реализаций, и что они заданы в координатах. Примерами таких реализаций могут быть печатные буквы и цифры.

Эталон также может передаваться в свободной реализации. Примерами могут служить буквы и цифры, написанные от руки. В этом случае, реализация формы в большей или меньшей степени похожа на эталон. Распознавание, то есть процесс отнесения такого визуального образа к своему эталону, зависит от присутствия в образе шума.

Предположим существование двух людей или предметов, один из которых посылает сигналы (он наывается источником и обозначен символом S), а второй – получающий эти сигналы (названный приемником и обозначенный символом R). Будем считать, что источник и приемник имеют в распоряжении одинаковые наборы эталонов. S

посылает эталон, пытаясь реализовать его в виде формы. R получает его в виде стимула и пытается воспринять и распознать этот стимул. Для интерпретации можно использовать следующую диаграмму:



Пометки (R) и (S) обозначают либо активность S или R , либо то, что результат относится к роли S или R .

7. Матрицы непрерывности

Как было зафиксировано в определении, форма является подпространством пространства E^2 . Дискриминант непрерывности формы зависит от способа рассмотрения нами формы в качестве комплекса. Теперь уникальным образом присвоим форме ее дискриминант разрывности (не принимаю во внимание множество возможностей рассматривать форму в качестве 1-комплекса). Возьмем, к примеру, два эталона нуля – арабский и в виде многоугольника.



$$\tau = 1(2)1$$



$$\tau = 1(2)2(2)1$$



$$\tau = 1(2)2(2)1$$

Во втором случае жирные точки являются концевыми точками (гладких) дуг многоугольника, а в третьем – одна точка является лишней. Во всех трех случаях мы имеем одно и то же число (две) дуг на интервале $[0, a]$. Склонимся также к тому, чтобы сжать $(2)2(2)$ до (2) .

Пусть дана форма F , а $[0, a]$ – ее область определения. Точка $t \in (0, a)$ – *регулярна*, если существует $\varepsilon > 0$, такое что многозначная

функция F непрерывна на $(t' - \varepsilon, t' + \varepsilon)$. Иначе точку $t' \in [0, a]$ назовем *сингулярной*.

Пусть

$$t_0 = 0 < t_1 < \dots < t_k = a$$

– последовательность всех сингулярных точек формы F . Теперь мы можем сопоставить форме следующую последовательность:

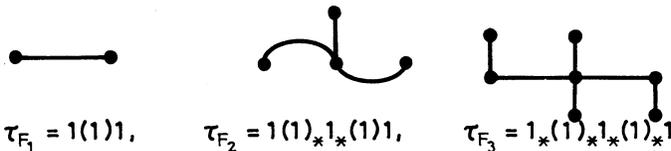
$$\tau_F = \tilde{n}_{t_0}(n_t^{(0)})\tilde{n}_{t_1} \dots \tilde{n}_{t_{k-1}}(n_t^{(k-1)})n_{t_k}$$

точно также, как мы делали это в разделе 4. Эта последовательность τ_F будет нами называться *каноническим дискриминантом непрерывности F* .

Легко видеть, что каждый дискриминант непрерывности F может быть приведен к каноническому дискриминанту непрерывности путем заменой термов вида $(m)m(m)$ на (m) в конечном числе случаев. Поскольку $-m + m - m = m$, мы видим, что эта замена не меняет значения выражения в левой части формулы в утверждении 3. Таким образом, как было доказано, канонический τ_F связан с Эйлеровой характеристикой формы.

Когда в разделе 4 вводились (линейные) формы, мы предположили, что они являются связными пространствами. Если теперь отказаться от этого предположения, лежащее в основе формы пространство можно считать объединением непересекающихся форм. Все сформулированные определения и утверждения справедливы и для этих более общих объектов. Рассмотрение таких объектов не находится в соответствии с нашим стремлением к обобщению, но процедура, которую мы собираемся описать, заставит нас сделать это обобщение. Грубо говоря, эта процедура будет декомпозицией формы на простые формы.

Простая форма – это такая форма F , такая что для каждой точки $t \in [0, a]$, $n_t = 1$. Это является примерами таких форм



Мы видим, что простая форма включает также в себя концевые компоненты.

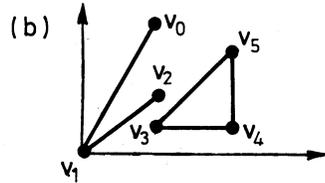
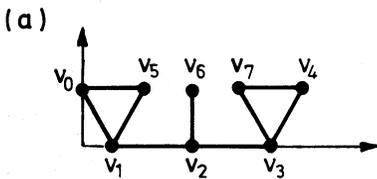
Пусть дано объединение непересекающихся форм S . Нас будет интересовать выбор тех простых форм L :

(I) которые являются *подформами* S , то есть объединениями дуг формы S ,

(II) которые являются *максимальными*, то есть такими, к которым нельзя добавить дуги.

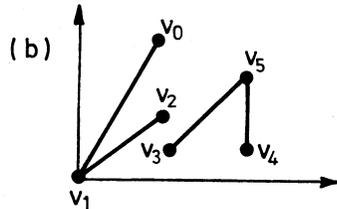
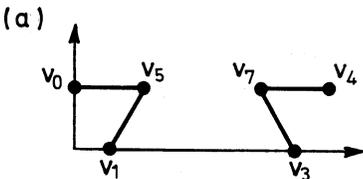
(III) которые являются *нижними* в форме S , то есть для каждой точки $(x, y) \in L$ и $(x, y') \in S$ L выполнено $y' > y$. Такие подформы L будем называть *слоями* формы S .

Рассмотрим некоторые примеры.



В случае (a) слоем является объединение следующих дуг: $\nu_0\nu_1, \nu_1\nu_2, \nu_2\nu_6, \nu_2\nu_3, \nu_3\nu_4$, а в случае (b) слой состоит из следующих дуг: $\nu_3\nu_4, \nu_4\nu_5$.

Удаляя внутренние интервалы горизонтальных дуг в L , а затем те вертикальные компоненты, которые являются изолированными в L , мы получим

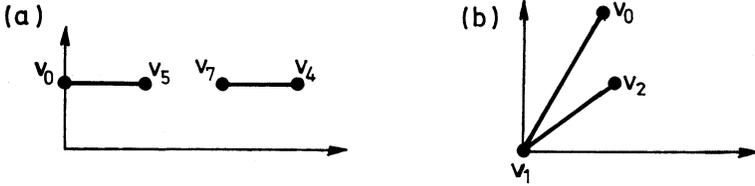


В обоих случаях полученные таким образом пространства являются разрывными (что и рассматривается в качестве мотивации для введения в рассмотрения объединений форм S).

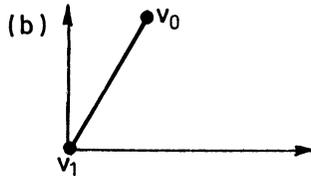
Теперь в случае (a) мы имеем два слоя: $\nu_1\nu_5$ и $\nu_3\nu_7$. В случае (b)

слоем является объединение таких дуг: $\nu_3\nu_5, \nu_5\nu_4$.

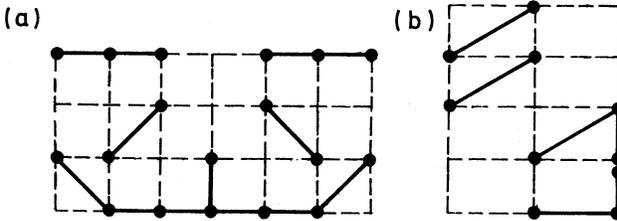
Проведя процедуру удаления снова, получим:



и в случае 2, после очередного удаления



Декомпозицией на слои будет:



Сопоставим каждой декомпозиции матрицу с числом строк, соответствующим числу слоев, а каждый ряд сопоставим дискриминанту непрерывности соответствующего ему слоя. Уже напечатанный конечные компоненты писать не будем, заменяя их на 0.

(a)

$$\begin{bmatrix} 0 & (1) & 1 & (1) & 0 & (0) & 0 & (0) & 0 & (1) & 1 & (1) & 0 \\ 0 & (0) & 0 & (1) & 1 & (0) & 0 & (0) & 1 & (1) & 0 & (0) & 0 \\ 1 & (1) & 1 & (1) & 1 & (1) & 1 & (1) & 1 & (1) & 1 & (1) & 1 \end{bmatrix}$$

$$\tau_{s_1} = 1(2)2(3)2_*(1)_*1_*(1)_*2(3)2(2)1$$

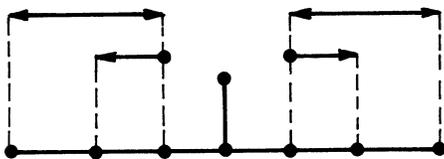
(b)

$$\begin{bmatrix} 0 & (1) & 1 & (0) & 0 \\ 1 & (1) & 1 & (0) & 0 \\ 0 & (0) & 0 & (1) & 0 \\ 0 & (0) & 1 & (1)_* & 1 \end{bmatrix}$$

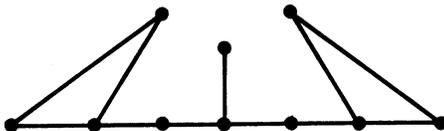
$$\tau_{s_2} = 1(2)_* 3_*(2)_* 1$$

Заметим, что сумма элементов в столбце равна соответствующему элементу τ_s .

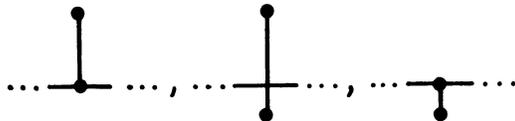
Заметим также, что по такой матрице мы можем восстановить форму, сохраняя ее существенные свойства. Например, в случае (a) строки геометрически представляют из себя следующее:



и соединяя эти части, получаем



что является полигональной реализацией. Компонента в центральной точке остается неопределенной, так что любой из следующих трех случаев возможен:



Такое геометрическое представление рядов является декомпозицией формы на несвязные части (что не является плохим случаем при декомпозиции на слои).

Теперь мы должны в целом установить способ проведения подобных декомпозиций, а также способ сопоставления форме матрицы непрерывности. Во-первых, начнем со следующего утверждения.

Утверждение 4. Любое объединение непересекающихся форм S состоит как минимум из одного слоя.

Доказательство. Посмотрим на S как на график многозначной функции. Тогда функция S определена на конечном числе непересекающихся замкнутых интервалов. Пусть первый из них — $[a, a_1]$. Пусть

$$t_0 = 0 < t_1 < \dots < t_k = a_1$$

является последовательностью всех сингулярных точек формы S , содержащихся в $[0, a_1]$. (Сингулярной точкой S будем называть то же, что мы подразумевали под этим в случае точки.)

Рассмотрим семейство $\{\alpha_i\}$ нижних дуг на каждом из интервалов $[t_{i-1}, t_i]$, $i = 1, \dots, k$. Для каждого $i \geq 2$ отметим у дуги α_i левую концевую компоненту, если правая концевая компонента α_{i-1} расположена выше левой концевой у α_i . Будем считать также, что левая концевая компонента у α_1 — отмечена.

Пусть s — максимальное из чисел $\{1, 2, \dots, k\}$, такое что у дуги α_s отмечена левая и правая компоненты.

Если $s = k$, то α_s со своими концевыми компонентами является слоем.

Предположим, что $s < k$. Тогда у дуги α_{s+1} левая компонента не является отмеченной. Если эта компонента выше правой концевой компоненты α_s , то α_s со своими концевыми компонентами является снова слоем. Если же правая концевая компонента α_s совпадает с левой концевой компонентой у α_{s+1} , то объединение дуг α_s и α_{s+1} вместе со своими концевыми компонентами является простой формой.

Повторим теперь то же рассуждение в случае точки t_{s+1} и дуги α_{s+2} . В этом случае α_{s+1} вместе со своими концевыми компонентами либо являются слоем, либо мы продлеваем эту простую форму, добавляя α_{s+2} .

За конечное число шагов мы получим слой формы S .

Пусть дана форма F , $[0, a]$ — ее область определения, а

$$0 = t_0 < t_1 < \dots < t_k = a$$

является последовательностью ее сингулярных точек. Семейство всех слоев F является конечным объединением простых форм. Обозначим это объединение через Λ_1 . Пусть τ_{Λ_1} является последовательностью, состоящей из элементов вида \tilde{n}_{t_i} и $(n_t^{(i)})$, как это определено в разделе 4, считая, что $\tilde{n}_{t_i} = 0$, когда прямая $x = t_i$ не пересекает Λ_1 и $n_t^{(i)} = 0$, когда в Λ_1 не дуги над отрезком $[t_i, t_{i+1}]$.

Чтобы упростить наши описания, будем называть дугу формы F *растяжкой*, если она проецируется на некоторый отрезок $[t_i, t_{i+1}]$ (с проекцией $(x, y) - x$). Все остальные дуги F принадлежат вертикальным прямым и мы будем называть их *вертикальными*. Эти дуги являются концевыми компонентами некоторых «растяжек».

Удаляя внутренние части всех растяжек в Λ_1 , а также концевые компоненты, не являющиеся концевыми компонентами каких-либо из оставшихся дуг, мы получим подмножество S_1 множества F , являющееся объединением непересекающихся форм (то есть, для которого уже не требуется связность).

Согласно утверждению 4, семейство всех слоев S_1 не является пустым. Обозначим это семейство через Λ_2 .

Пусть τ_{Λ_2} является последовательностью, определенной также, как для случая Λ_1 с тем исключением, что $\tilde{n}_{t_i} = 0$, если эта компонента была уже включена в τ_{Λ_1} .

Снова удаляя внутренние части всех растяжек в Λ_2 вместе с концевыми компонентами, не являющимися концевыми компонентами каких-либо из оставшихся дуг, мы получим $S_2 \subset S_1$. Обозначим теперь через Λ_3 семейство всех слоев S_2 . Пусть τ_{Λ_3} является соответствующей последовательностью с $\tilde{n}_{t_i} = 0$, если эта компонента была уже включена в $\tau_{\Lambda_1} \cup \tau_{\Lambda_2}$.

Продолжая этот процесс по индукции, мы на каждом шаге будем удалять как минимум одну растяжку. Следовательно, число шагов конечно, обозначим его через m .

Объединяя эти последовательности в виде матрицы, принимая во внимание последовательность слоев, мы получим:

$$\begin{bmatrix} \tilde{\tau}_{\Lambda_m} \\ \dots \\ \tilde{\tau}_{\Lambda_2} \\ \tilde{\tau}_{\Lambda_1} \end{bmatrix}$$

Эту матрицу назовем *матрицей непрерывности* формы F .

Набор слов $\Lambda_1, \Lambda_2, \dots, \Lambda_m$ будет называться *декомпозицией* формы F на слов.

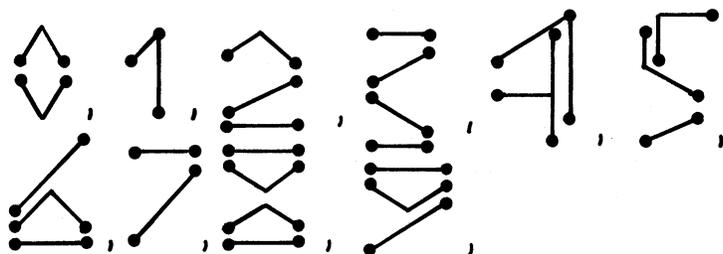
Матрица простой формы, состоящая из одной строки, может рассматриваться как дискриминант непрерывности.

8. Матрицы непрерывности для цифр

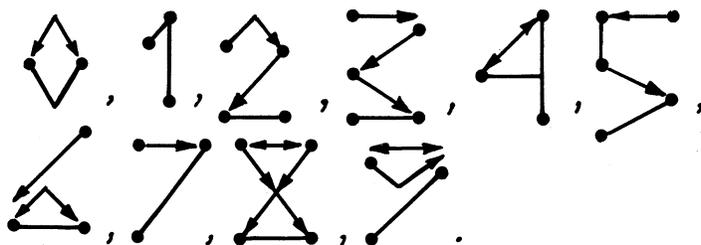
Теперь снова вернемся к примеру с цифрами как формами, заданными своими эталонами (см. раздел 5). Каждой цифре сопоставим матрицу непрерывности

 $\begin{bmatrix} 0(1)0 \\ 1(1)1 \\ 1(2)1 \end{bmatrix}$	 $\begin{bmatrix} 1(1)_*1 \\ 1(1)_*1 \end{bmatrix}$	 $\begin{bmatrix} 1(1)0 \\ 0(1)1 \\ 1(1)1 \\ 2(3)2 \end{bmatrix}$
 $\begin{bmatrix} 1(1)0 \\ 0(1)1 \\ 1(1)0 \\ 1(1)1 \\ 3(4)2 \end{bmatrix}$	 $\begin{bmatrix} 0(1)0 \\ 1(1)_*1 \\ 1(2)_*1 \end{bmatrix}$	 $\begin{bmatrix} 0(1)1 \\ 1_*(1)0 \\ 1(1)0 \\ 2_*(3)2 \end{bmatrix}$
 $\begin{bmatrix} 0(1)1 \\ 0(1)0 \\ 1(1)1 \\ 1(3)2 \end{bmatrix}$	 $\begin{bmatrix} 1(1)0 \\ 1(1)1 \\ 2(2)1 \end{bmatrix}$	 $\begin{bmatrix} 0(1)1(1)0 \\ 1(1)0(1)1 \\ 0(1)1(1)0 \\ 1(1)1(1)1 \\ 2(4)3(4)2 \end{bmatrix}$
	 $\begin{bmatrix} 0(1)0 \\ 1(1)0 \\ 1(1)1 \\ 2(3)1 \end{bmatrix}$	

Декомпозиции на слои этих форм приведены ниже:



а декомпозиция на непересекающиеся части, соответствующие строкам матрицы непрерывности приведены ниже:



Мы можем видеть, что все канонические дискриминанты непрерывности являются последовательностями числа, за исключением случая цифры «8», когда последовательность состоит из 4 членов. Мы также видим, что различные эталоны цифр имеют различные дискриминанты непрерывности, а также разные матрицы непрерывности.

«Плохая» реализация обычно воспринимается как форма с некоторым количеством шума, добавленного к соответствующему эталону. Говоря более конкретно, рассмотрим пример цифры «2». Здесь мы используем термины «сбалансированная» – «несбалансированная», «симметричная» – «асимметричная» (нерегулярная), «полная» – «неполная», описывая их значения в случае выбранного примера.

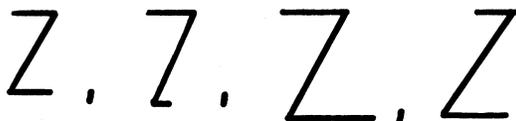
Хорошо сбалансированная цифра «2» имеет следующую матрицу непрерывности:

$$\begin{bmatrix} 1 & (1) & 0 \\ 0 & (1) & 1 \\ 1 & (1) & 1 \end{bmatrix},$$

в которой сингулярные точки $t_0 = 0 < t_1 = a$, а все слои являются интервалами. Самой симметричной реализацией является следующая фигура:



Во-первых, добавим шум к этой фигуре, чтобы привести к ее разбалансированности. Как результат, получим следующую последовательность форм:



(нет шума)



Матрицы непрерывности этих форм все различны, а для последней формы матрица имеет вид

$$\begin{bmatrix} 0 & (0) & 0 & (0) & 1 & (1) & 0 \\ 0 & (1) & 1 & (1) & 1 & (1) & 1 \\ 1 & (1) & 1 & (0) & 1 & (1) & 0 \end{bmatrix}$$

Во-вторых, добавим к каждой из этих форм шум, чтобы привести к асимметрии, заменяя дуги случайными (кусочно-непрерывными) гладкими дугами, но в каждом случае сохраняя матрицы непрерывности. После этой процедуры мы получим:



Этим способом может быть получено большое разнообразие форм, и некоторые из них могут быть легко сопоставлены образам других цифр. Но распознавание подразумевает набор эталонов и, в этом случае, все эти формы имеют символическое значение цифр «2».

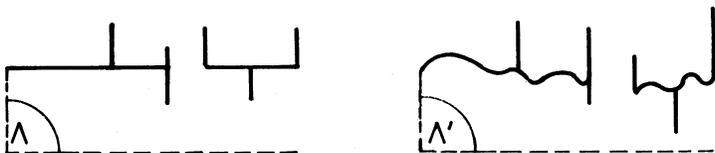
Это значение основано на общей глобальной структуре этих форм которая состоит из последовательности трех слоев, каждый из которых является интервалом и связан с остальными способом, представленным каждой из матриц непрерывности (в простом случае это матрица

$$\begin{bmatrix} 1 & (1) & 0 \\ 0 & (1) & 1 \\ 1 & (1) & 1 \end{bmatrix}$$

9. Эквивалентность форм

Классификация форм по их топологическим типам была бы слишком громоздкой. Поэтому мы должны подумать о классификации, сохраняющей некоторые комбинаторные и позиционные свойства форм.

Семейство слоев, рассматриваемое независимо от формы, которой оно принадлежит, является последовательностью отдельных простых форм. Теперь, в более строгой форме, скажем, что два семейства слоев являются гомеоморфными, если существует гомеоморфизм, сохраняющий порядок их дуг, вертикальность компонент и их ориентацию. например, следующие два семейства слоев:



являются гомеоморфными, и они не гомеоморфны ни одному из следующих двух семейств:



(и $\Lambda'' \neq \Lambda'''$). Зафиксируем это значение гомеоморфизма в виде определения.

Два семейства слоев Λ и Λ' являются *гомеоморфными*, если существует гомеоморфизм $h : \Lambda \rightarrow \Lambda'$, такой что $h(x, y) = (h^1(x), h_x^2(y))$, где h^1 является возрастающей функцией по x и для каждого x h_x^2 является возрастающей функцией по y . В этом случае мы говорим, что h является *гомеоморфизмом семейств слоев*.

Каждый способ распознавания предполагает как минимум эмпирический способ классификации форм.

Способ, которому мы следуем, предполагает наличие классов форм, заданных эталоном, который варьируется путем искажения глобальным шумом: геометрическая нерегулярность и несбалансированность дуг. Глобальная структура формы, к которой относится наше определение эквивалентности, является той же декомпозицией в гомеоморфных семействах слоев.

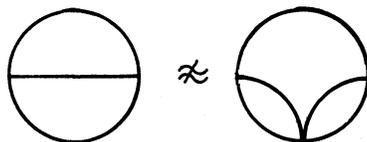
Две формы F и F' являются *эквивалентными*, если в результате декомпозиции они распадаются на одно и то же число семейства слоев

$$\Lambda_1, \dots, \Lambda_n \text{ и } \Lambda'_1, \dots, \Lambda'_n,$$

соответственно, и существует гомеоморфизм $h : F \rightarrow F'$, такой что для каждого i ($i = 1, \dots, n$) ограничение $h|_{\Lambda_i}$ является гомеоморфизмом семейств Λ_i и Λ'_i . В этом случае мы называем отображение h *гомеоморфизмом форм*.

Для двух эквивалентных форм будем говорить, что они имеют *один и тот же тип*.

Заметим, что две формы, имеющие две соответствующих семейства слоев гомеоморфными, не обязательно являются эквивалентными формами (и даже не обязательно имеют один и тот же топологический тип). Простым примером является следующая пара форм:



где гомеоморфизмы слоев $h_i : \Lambda_i \rightarrow \Lambda'_i$ не могут быть объединены так, чтобы они образовывали гомеоморфизм форм.

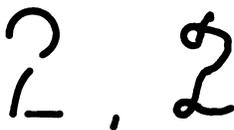
Заклучим настоящую работу некоторыми комментариями по работе и по нашим будущим идеям.

Некоторые комментарии

Эта работа отражает способ, с помощью которого были разработаны некоторые наши идеи. Мы намеренно оставили его в этом виде.

Мы предполагаем с достаточной степенью уверенности что некоторые из наших возможных заинтересованных читателей не будут знакомы с топологией. Некоторое знание в этой области, достаточное для нужд настоящей работы, может быть сформировано после прочтения разделов 2 и 3 (может также оказаться полезной помощь знакомого математика). Для более глубокого знакомства мы предлагаем следующие замечательные книги: [2], [4], [5].

Говоря с позиций человека R (раздел 6), никакие попытки классифицировать стимулы не могут быть признаны осмысленными. Две следующие реализации числа «2»:



могут рассматриваться как эталоны, искаженные локальным шумом, в данном случае являющимся неполнотой и самопересечением соответственно. Они не принадлежат к одному классу эквивалентности форм, определенных эталоном «2». Способы, с помощью которых они могут быть восприняты как типы форм, заданные «2», формально представляются как стрелка, озаглавленная словом «восприятие» на треугольной диаграмме в разделе 6. Мы также планируем заняться изучением подобных способов.

Матрица непрерывности показывает глобальную структуру формы. Когда две эквивалентные формы имеют разные матрицы, они имеют различные степени разбалансированности (чем больше столбцов в матрице, тем больше разбалансированность). Кажется интересным попытаться выразить понятие эквивалентности форм в терминах их матриц непрерывности как минимум для некоторых подклассов форм, что мы также планируем сделать. В практическом смысле эти матрицы могут уже эффективно использоваться в слу-

чае некоторых специфических наборов эталонов.

В компинаторном смысле форма лучше связана, если каждое семейство Λ_i состоит из одного элемента. Подкласс таких форм кажется довольно большим, и в этом подклассе рассмотрение некоторых объектов упрощается.

В общем, мы можем сказать, что чем больше классы эквивалентности, тем более эффективно распознавание. Но, например, если набор эталонов включает в себя число «2» и букву «Z», тогда они имеют один и тот же тип. Если посмотреть на форму с позиции оси y , все происходит аналогично и мы имеем дискриминант непрерывности τ_F^y и матрицу непрерывности, составленную из вертикальных элементов в качестве слоев. С этой точки зрения, число «2» и буква «Z» различны.

Заметим, что в качестве логического упрощения, предположение о гладкости может быть заменено непрерывностью во всех рассмотрениях, в которых это предположение имело место.

Существует большое количество работ, посвященных проблеме распознавания образов. Не претендуя на полноту, мы приводим в списке литературы только некоторые из них, в которых рассмотрен топологический подход. Когда мы сравнивали общие подходы, мы нашли, что наша работа ближе всего находится к работе Н. Nishida и S. Mori.

Список литературы

- [1] Kolher W. Gestalt Psychology. New York: Liveright, 1947.
- [2] Kuratowski K. Topology. I, II. Academic Press, 1966, 1968.
- [3] Mantas J. An Overview of Character Recognition Methodologies // Pattern Recognition. Vol. 19. 1986. P. 425–430.
- [4] Maunder C.R.F. Algebraic Topology. London: Van Nostrand, 1970.
- [5] Munkres J.R. Elements of Algebraic Topology. Addison-Wesley, 1984.
- [6] Nishida H., Mori S. Algebraic Description of Curves Structure // IEEE trans. Patt. Anal. Machine In. Vol. 14. No. 5. May 1992. P. 516–533.

- [7] Ramesh S.R. A Generalized Character Recognition Algorithm: A Graphical Approach // Pattern Recognition. Vol. 32. No. 4. 1989. P. 347–350.
- [8] Shridhar M., Badreldin A. Recognition of Isolated and Simply Connected Handwritten Numerals // Pattern Recognition. Vol. 19. No. 1. 1986. P. 1–12.
- [9] Wertheimer M. Productive Thinking. New York: Harper and Brothers, 1945.