# Билет 8. Корреляционные методы сопоставления изображений. Метод фазовой корреляции

*Амплитудно-фазовая корреляция*

Проблема сопоставления изображений может решаться с привлечением разных типов представлений, начиная от математических и заканчивая представлениями знаний. Классическими методами сопоставления являются корреляционные методы, базовая идея которых крайне проста и заключается в попиксельном сравнении изображений, хотя различия в реализациях могут быть достаточно значительны.

В простейшем случае мера сходства двух изображений как функция взаимного смещения  вычисляется как квадрат среднеквадратичного отклонения яркостей соответствующих (с учетом величины смещения) пикселей

, (2.10)

где *f*1 и *f*2 – изображения одного и того же объекта, отличающиеся только сдвигом; *N*x*N* – размер области, в которой вычисляется корреляция.

Как видно из формулы (2.10), данный подход позволяет определять взаимные смещения двух изображений при отсутствии других геометрических преобразований, а также яркостных изменений.

Величину  несложно преобразовать к виду



Полагая величины  и  не зависящими от смещения (что верно при некоторых оговорках), можно получить, что минимизация критерия СКО (2.10) эквивалентна максимизации значения корреляции:

. (2.11)

Несложно убедиться, что положение корреляционного максимума при линейном глобальном преобразовании яркости  не меняется (только при  максимум превращается в минимум). Таким образом, данный подход позволяет сопоставлять изображения при линейных глобальных искажениях яркости.

Если смещения  и  пробегают значения от –*N* до *N*, то несложно убедиться, что число операций при вычислении кросскорреляционного поля по формулам (2.10) или (2.11) будет пропорционально *N*4.

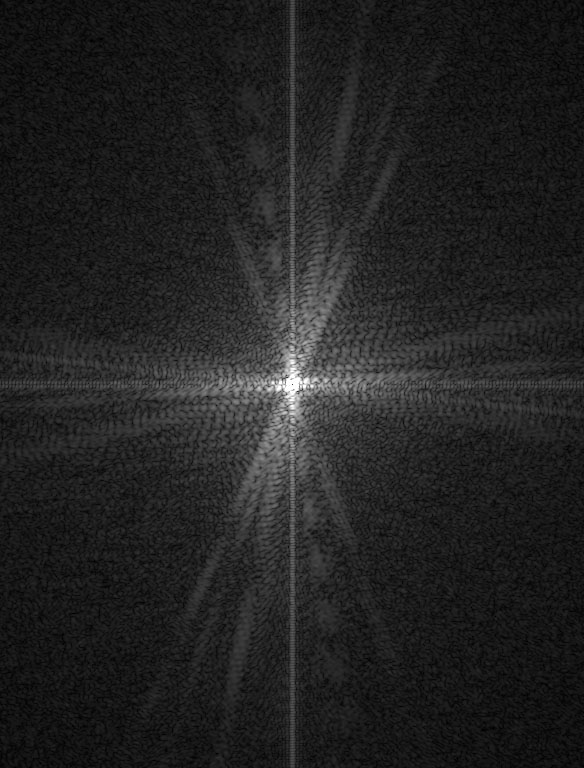
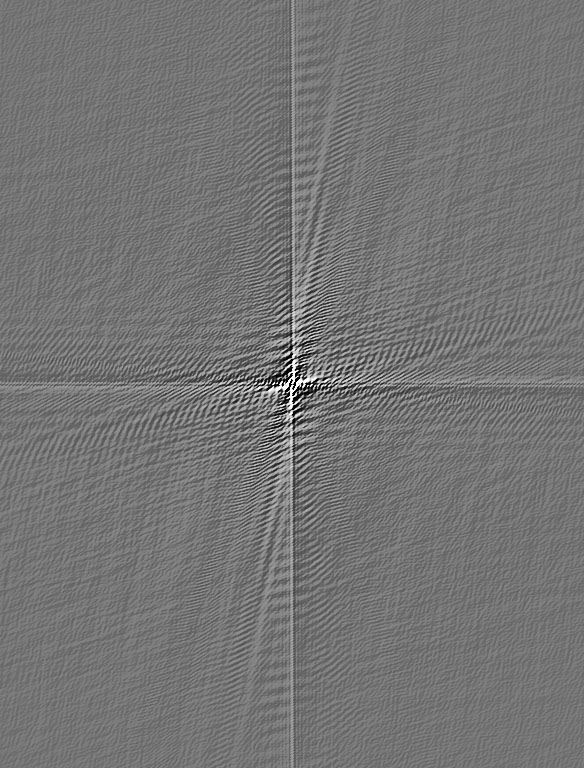
Однако существует возможность построения кросскорреляционного поля  за меньшее число операций, если воспользоваться быстрым преобразованием Фурье и тем фактом, что Фурье-образ кросскорреляционного поля может быть получен как произведение спектров изображений.

По определению, прямое и обратное дискретные преобразования Фурье, ДПФ, (см. рис. 2.10) вычисляются как

 (2.12)

где *u* и *v* – пространственные частоты. При вычислении пространственного спектра по этим формулам число операций также пропорционально *N*4, однако алгоритм быстрого преобразования Фурье позволяет произвести вычисление за время, пропорциональное .

Рис. 2.10. Пример вычисления преобразования Фурье: (а) исходное изображение; (б) вещественная компонента спектра; (в) амплитудный (энергетический) спектр



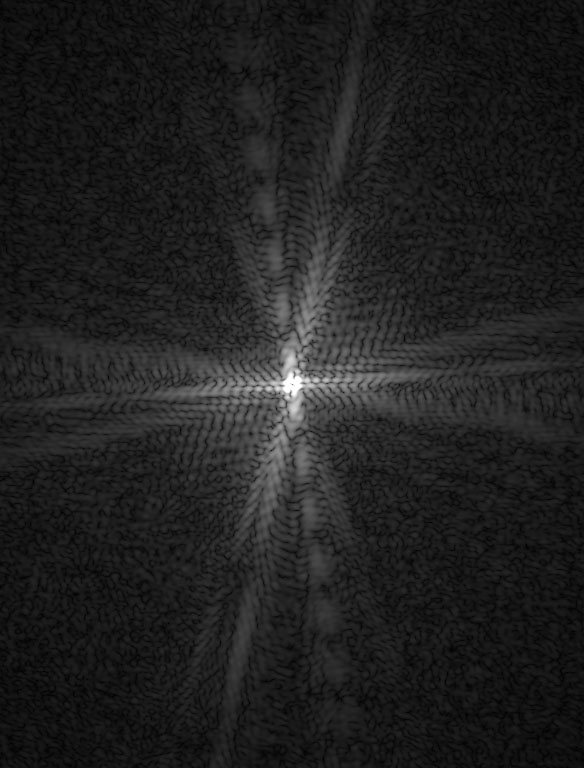
Кросскорреляционное поле тогда может быть получено как

 (2.13)

что также может быть вычислено за время, пропорциональное  ( – комплексное сопряжение). Таким образом, общее время работы при вычислении кросскорреляционного поля через ДПФ будет пропорционально , хотя коэффициент пропорциональности значительно больше, чем при непосредственном вычислении по формуле (2.11), в связи с чем данный подход нерационально использовать при малых *N*.

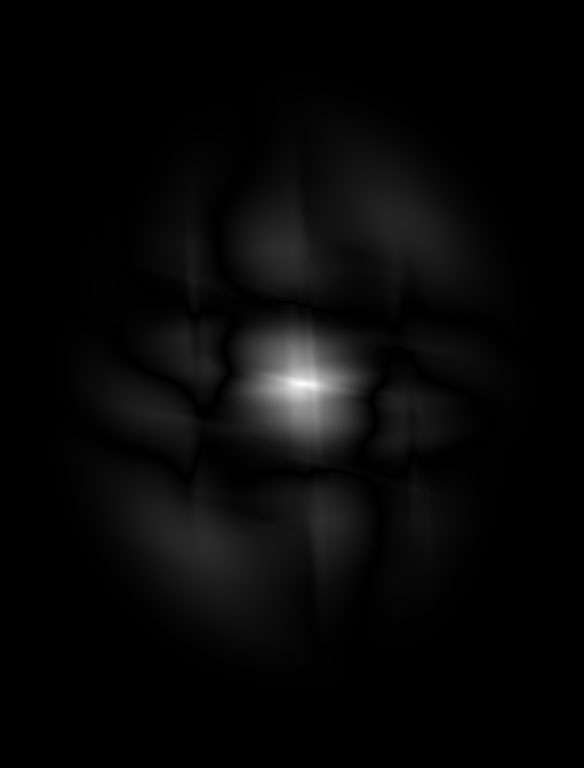
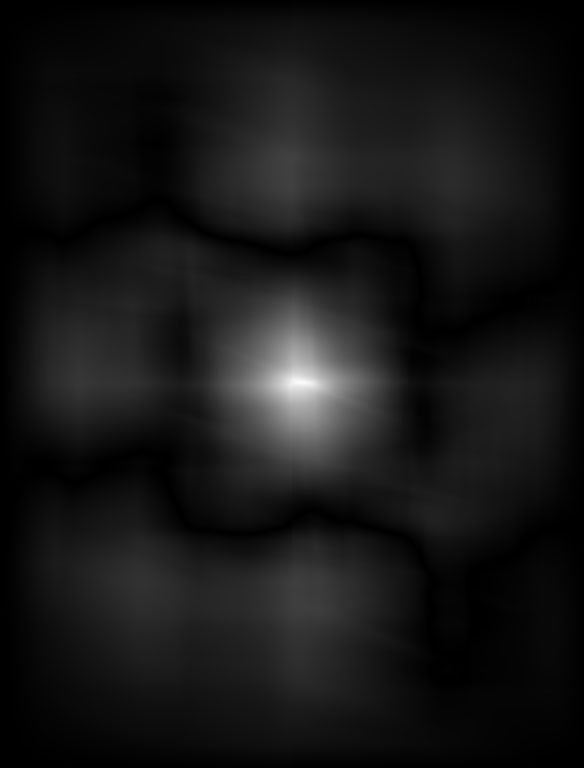
Необходимо отметить, что в алгоритмах ДПФ изображение трактуется как периодический сигнал, то есть копии изображения как бы приставлены друг к другу, из-за чего появляется резкий переход от одного края изображения к другому. Это вызывает появление «креста» в центре спектра изображения (см. рис. 2.10) и, как мы увидим ниже, ухудшение корреляционного отклика. Плавное приведение интенсивностей у краев к среднему значению устраняет эффект края. Как видно из рис. 2.11, центральный «крест», который отчетливо виден на рис. 2.10, пропадает. При этом также рекомендуется расширять изображение, помещая его на поле однородной яркости (тогда коррелируемое с ним изображение не будет попадать на соседние периоды).

Рис. 2.11. Изображение, усредненное к краям, и его амплитудный спектр



На рис. 2.12 представлен пример вычисления кросскорреляционного поля без устранения эффекта края и с его устранением. На рисунке приведены два изображения, полученные с разных ракурсов. Поскольку это изображения трехмерного объекта, а проекция центральная, то геометрическое преобразование изображения не описывается одним общим смещением (но сдвиги для всех точек достаточно близки), в связи с чем максимум на кросскорреляционном поле не выражен четко. Однако видно, что максимум может быть локализован, причем при использовании сглаживания к краю кросскорреляционное поля выглядит контрастнее, и максимум может быть определен точнее. В действительности, эффект края приводит к появлению ложного максимума в центре поля (точке, соответствующей смещению (0, 0)) и к размытию максимума, соответствующего истинному смещению.

Кроме того, из рис. 2.12 видно, что значение максимума на кросскорреляционном поле не сильно отличается от его окружения. Существуют методы, позволяющие осуществлять обнаружение и локализацию максимума с более высокой точностью.



а)

б)

в)

г)

Рис. 2.12. (а) и (б) пара сопоставляемых изображений, снятых с разных ракурсов; (в) кросскорреляционное поле двух изображений, вычисленное с помощью ДПФ; (г) кросскорреляционное поле изображений со сглаженными краями

*Фазовая корреляция*

Корреляционные методы, описанные выше, называются амплитудно-фазовыми, поскольку в них используется как амплитудная, так и фазовая спектральная информация. Помимо этого используется фазовая корреляция, при вычислении которой амплитудная информация не используется:

. (2.14)

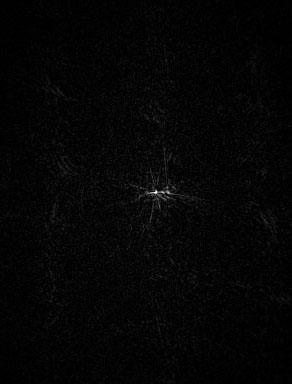
Напомним, что Фурье-образы  и  являются комплекснозначными, даже если исходные изображения представлены вещественнозначными функциями. Значения  и  являются модулями соответствующих комплексных величин.

Особенность фазовой корреляции заключается в том, что для нее автокорреляционное поле некоторой функции представляет собой дельта-функцию, в чем несложно убедиться

. (2.15)

На рис. 2.13 представлено кросскорреляционное поле (в случае фазовой корреляции) изображений, приведенных на рис. 2.12а,б. Видно, что здесь максимум может быть локализован с большей точностью (при этом видно, что в данном случае выделяется несколько максимумов, соответствующих сдвигу объекта на переднем плане и фона, имеющих разные смещения).

Рис. 2.13. Пример кросскорреляционного поля, получающийся в случае фазовой корреляции вместо поля, представленного на рис. 2.12г.



Следует отметить, что в случае фазовой корреляции также полезным является сглаживание изображений к краю с дальнейшим расширением области изображений с заполнением области расширения средним значением (см. рис. 2.14).

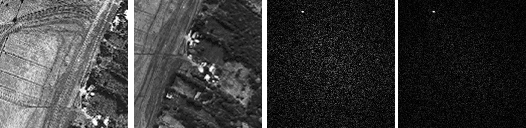
Рис. 2.14. Пример уменьшения шумов на кросскорреляционном поле в случае фазовой корреляции при устранении эффекта края: а) и б) исходные изображения; в) кросскорреляционное при использовании фазовой корреляции; г) кросскорреляционное поле при устранении эффекта края

а)

б)

в)

г)



Кроме лучшей локализации максимума метод фазовой корреляции в ряде случаев может также оказаться немногим менее чувствительным к нелинейным или локальным преобразованиям яркости, чем метод амплитудно-фазовой корреляции. Однако в этом подходе не используется большой объем информации, заключенный в амплитудной составляющей спектра, что может привести в ряде случаев и к понижению вероятности обнаружения максимума, соответствующего истинному смещению. Общим недостатком описанных методов является сильное ограничение на допустимое пространственное преобразование изображений, а именно, оно должно являться преобразованием сдвига.

# Дополнительно. Метод Фурье-Меллина, локальной корреляции и максимизации взаимной информации

*Метод Фурье-Меллина*

При вычислении корреляционного поля по формуле (2.11) нет принципиальных ограничений на то, чтобы варьировать только взаимное смещение двух изображений. В равной степени могут перебираться не только все возможные сдвиги, но также и углы поворота или, к примеру, взаимные масштабы изображений. Когда выбран конкретный набор параметров пространственного преобразования, после его выполнения может быть посчитана величина корреляции между одним изображением и преобразованным другим изображением. По максимуму корреляционного отклика можно таким образом определить параметры пространственного преобразования (при условии, что яркостные искажения не слишком отличаются от линейных).

При этом, однако, возникает трудность, связанная с вычислительной сложностью алгоритма сопоставления. Если полагать, что каждый параметр преобразования может принимать одно из *N* значений, то необходимое число операций будет пропорционально *Nm*+2, где *m* – число параметров. Поскольку уже для преобразования сдвига (при *m*=2) оказывается критичным повышение быстродействия алгоритма за счет использования БПФ, то при большем числе параметров время сопоставления оказывается неприемлемым.

Однако вместо поиска максимума во всем пространстве параметров взаимного геометрического преобразования изображений существует возможность так преобразовать изображения, чтобы результат преобразования оказался инвариантным к некоторой подгруппе исходной группы геометрического преобразования.

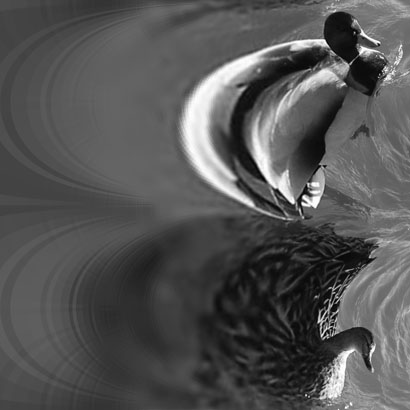
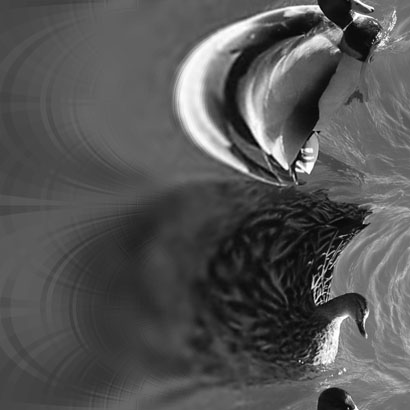
В частности, амплитудный спектр изображения  остается постоянным при сдвиге этого изображения (меняется лишь фаза гармоник пространственного спектра, но не их амплитуда). Поскольку помимо взаимных сдвигов изображения одного и того же объекта часто подвержены вращению и масштабированию, возникает вопрос, как преобразуется спектр при масштабировании и вращении изображения. Несложно убедиться, что масштабирование и вращение исходного изображения приводит к такому же масштабированию и вращению его спектра. Таким образом, амплитудный спектр изображения  инвариантен к сдвигу изображения, но сохраняет тот же масштаб и угол поворота.

Можно было бы рассматривать амплитудные спектры  и  в качестве изображений, которые необходимо сопоставить при условии, что они отличаются только масштабом и углом поворота. При этом, казалось бы, нужно перебирать разные значения этих параметров пространственного преобразования и вычислять значение корреляции по формуле (2.11). Однако и здесь можно воспользоваться быстрым вычислением корреляционного поля через БПФ, если перейти в такую систему координат, в которой масштаб и поворот превращаются в сдвиги по двум осям. Такой системой координат является полярно-логарифмическая система координат, переход к которой определяется соотношениями:

. (2.16)

На рис. 2.15 представлен пример перехода к полярно-логарифмическим координатам для изображений, отличающихся масштабом и углом поворота. Из рисунка видно, что изменение масштаба и угла поворота действительно приводит к сдвигу изображения в данной системе координат (напомним, однако, что данное преобразование координат подразумевается применять не к исходным изображениям, которые могут быть смещены друг относительно друга, а их амплитудным спектрам).

Рис. 2.15. Пример перехода в полярно-логарифмическую систему координат при масштабировании и вращении изображения



Поворот,

масштаб

Сдвиг

по X и Y

Полярно-логарифмическое преобразование

В итоге, после отображения амплитудного спектра изображения в полярно-логарифмическую систему координат значения взаимного масштаба и вращения могут быть найдены как сдвиги, для чего может быть использован обычный корреляционный метод с вычислением кросскорреляционного поля с помощью БПФ. После нахождения масштаба и угла поворота они могут быть компенсированы для исходных изображений, после чего уже может быть найден и взаимный сдвиг.

Таким образом, мы приходим к методу Фурье-Меллина, позволяющему сопоставлять изображения при взаимном преобразовании из группы подобия за время порядка . Этот метод состоит из следующих шагов

1. Для данных изображений  и  построить их амплитудные спектры  и .
2. Перевести амплитудные спектры в полярно-логарифмическую систему координат:

.

1. Построить кросскорреляционное поле «изображений»  и  в плоскости параметров  и найти на нем максимум .
2. Компенсировать взаимный масштаб  и угол вращения  для исходных изображений, после чего построить кросскорреляционной поле и найти на нем максимум , который будет соответствовать взаимному сдвигу.

Таким образом, будут найдены все параметры преобразования группы подобия.

На рис. 2.16 представлен пример успешного совмещения пары изображений методом Фурье-Меллина. Как видно из рисунка, изображения отличаются не только сдвигом, но и углом поворота.

Хотя метод Фурье-Меллина позволяет расширить класс допустимых взаимных пространственных преобразований, при этом несколько снижается надежность определения параметров преобразования, поскольку на каждом шаге данного метода происходит потеря части информации (при построении амплитудного спектра причина этой потери очевидна, однако заметные потери появляются и при переходе к полярно-логарифмической системе координат из-за дискретности координатной сетки).

Рис. 2.16. пример совмещения изображений методом Фурье-Меллина



Несмотря на расширение класса допустимых пространственных преобразований в методе Фурье-Меллина, это решение оказывается недостаточно общим, поскольку оно не подходит как для аффинных или проективных глобальных преобразований, так и для локальных преобразований, которые часто встречаются в реальных задачах компьютерного зрения. В рамках корреляционного подхода существует возможность дальнейшего расширения класса допустимых пространственных преобразований.

*Метод локальной корреляции*

Многие методы сопоставления изображений работают с ограниченными моделями пространственных преобразований (например, группы подобия или аффинной группы). При этом преобразования выбранного являются хотя и не точной, но неплохой аппроксимацией реального взаимного преобразования изображений. В результате, эти методы оказываются относительно работоспособными, но не точными. После компенсации основных компонентов пространственного преобразования остаются некомпенсированными компоненты следующего порядка малости.

К примеру, изображения, представленные на рис. 2.12а,б обладают некоторым взаимным смещением, после компенсации которого остаются небольшие собственные смещения для каждой точки изображений. Также, например, и для некоторых аэрокосмических изображений после компенсации аффинного преобразования может остаться незначительная проективная компонента, которая может быть описана в терминах локальных сдвигов, приписанных каждому пикселю.

Сходная ситуация имеет место для изображений биологических объектов, например, изображений, полученных в результате томографии головного мозга. В целом изображения могут быть пространственно совмещены (если компенсирован поворот головы относительно томографа), но отдельные области мозга могут быть немного сдвинуты друг относительно друга.

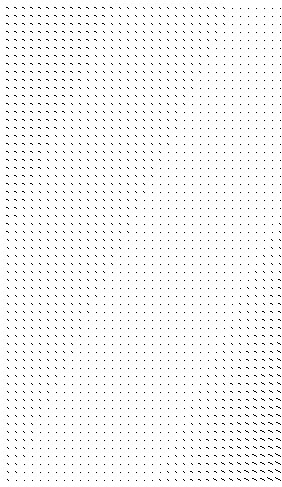
Если глобальное взаимное преобразование компенсировано, то окрестности сопряженных точек двух изображений должны отличаться незначительно. Это приводит к общей идее использовать корреляцию не для всей области изображений, а только для их фрагментов, взятых вокруг некоторых точек (см. рис. 2.17), что приводит к возможности построения поля локальных смещений (см. рис. 2.18).

Рис. 2.17. Пример вычисления локального смещения по фрагменту изображения в рамках корреляционного подхода



Локальное смещение

Рис. 2.18. Пример построения поля локальных смещений путем выполнения корреляции фрагментов, взятых вокруг набора точек, расположенных на регулярной сетке (белой линией отмечен овраг, по разные стороны от которого местность имеет разный наклон, что вызывает смещение соответствующих точек на изображениях, снятых с разных ракурсов)



Существует много вариантов реализации идеи локальной корреляции, однако мы не будем углубляться в технические подробности. Рассмотрим лишь пример использования здесь подхода с переменным разрешением. Поскольку при локальной корреляции берутся фрагменты фиксированного размера вокруг точек с одинаковыми координатами, то при локальном смещении, большем размера окна, локальное смещение в принципе не может быть определено. Однако если поле локальных смещений строить для изображений с пониженным разрешением, а далее их уточнять с постепенным повышением разрешения, то ограничение на допустимые локальные сдвиги будут слабее (рис. 2.19).

Рис. 2.19. Пример уточнения локального смещения для некоторой пары изображений



фрагменты

корреляционное поле

1/4x

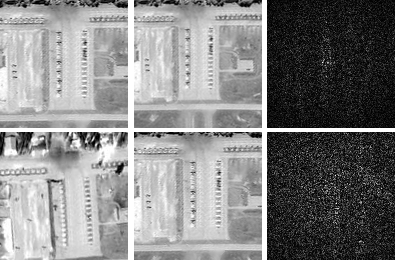
1/2x

1x

Подход с переменным разрешением позволяет расширить допустимые локальные сдвиги. Эти сдвиги могут появляться вследствие неполной компенсации взаимного глобального пространственного преобразования изображений. Однако если не полностью компенсированными остаются взаимный масштаб и поворот, то локальные фрагменты, взятые даже вокруг сопряженных точек (то есть с относительным нулевым смещением), будут подвержены тому же взаимному масштабу и повороту. Метод локальной корреляции, как и глобальный корреляционный метод, работающий только со сдвигами, весьма чувствителен к подобным геометрическим искажениям (см. рис. 2.20).

Как отмечалось ранее, корреляционные методы очень чувствительны к преобразованиям яркости, связанным с изменением условий освещения, сменой типов сенсоров и собственной изменчивостью объектов (например, сезонно-суточными изменениями для аэрокосмических изображений). Метод локальной корреляции менее чувствителен к подобным яркостным искажениям, поскольку допускает линейное преобразование яркости в каждом отдельном фрагменте изображения, а не только единое линейное преобразование яркости для всего изображения. Однако остается проблема компенсации глобального пространственного преобразования в условиях сложных яркостных искажениях.

Рис. 2.20. Пример разрушения корреляционного отклика при фазовой корреляции (а) фрагмента со своей копией, уменьшенной в 1.1 раза; (б) фрагментов двух разных изображений одной и той же местности при некомпенсированном угле поворота, равном 3o



а)

б)

*Совмещение изображений путем максимизации взаимной информации*

Можно считать, что корреляционные методы опираются на функциональные представления изображений. Можно обратиться и к более общим стохастическим моделям изображений, в которых яркости пикселей трактуются как отсчеты некоторой случайной величины. Тогда вероятности случайных величин, соответствующих двум совмещаемым изображениям, не будут статистически независимыми.

В случае использования простейшей стохастической модели яркости во всех точках изображения считаются реализациями независимых и одинаково распределенных случайных величин. Тогда для изображения *f*(*x*) как реализации некоторой случайной функции *F* может быть вычислена его энтропия

,

где *G* и *V* – области определения и значения , *E* – математическое ожидание случайной величины, *I* – количество информации, *P* – вероятность. В данной формуле  можно считать гистограммой изображения .

Условная энтропия двух изображений *F*1 и *F*2 примет форму

.

Взаимная информация, содержащаяся в изображениях, определяется как

. (2.17)

Для взаимной информации выполняются соотношения

 (2.18)

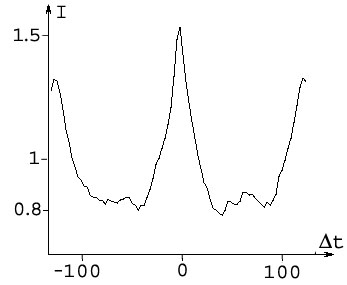
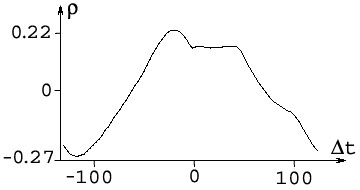
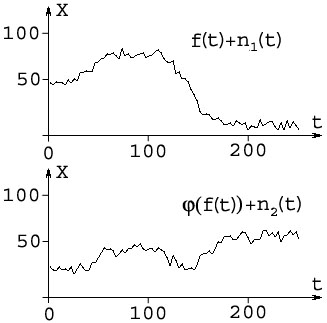
Кроме того, основным свойством взаимной информации является то, что ее значение для двух случайных величин инвариантно к произвольным функциональным преобразованиям этих величин и изменяется только при изменении степени их статистической независимости.

Иными словами, равенство нулю взаимной информации означает, что для любых невырожденных функций φ1 и φ2 выполняется равенство

. (2.19)

В то же время, равенство нулю коэффициента корреляции означает, что это равенство выполняется лишь для линейных преобразований φ1 и φ2. Таким образом, можно считать (в контексте сопоставления изображений), что критерий взаимной информации расширяет корреляционный критерий на случай нелинейных преобразований яркости (см. рис. 2.21).

Рис. 2.21: а) срезы двух изображений (с нулевым взаимным сдвигом) при нелинейных искажениях яркости; б) результат вычисления корреляции (максимум смещен); в) результат вычисления взаимной информации при разных сдвигах (выраженный максимум соответствует истинному смещению)



а)

б)

в)

Таким образом, вместо корреляционных методов могут использоваться методы совмещения изображений путем максимизации взаимной информации, в которых перебираются разные значения параметров пространственного преобразования, при которых вычисляется значение взаимной информации и находится максимум.

Несмотря на возможность работать с нелинейными преобразованиями яркости, в методах совмещения путем максимизации взаимной информации остается ограничение на то, что это преобразование должно быть глобальным (то есть единообразно применимым к любой точке изображения). Это в определенной степени верно, например, для случая смены сенсора, однако изменение условий освещения и собственная изменчивость объектов могут вызывать локальные, неоднородные, изменения яркости. То есть и здесь проблема инвариантности яркостным преобразованиям не решается. Кроме того, для методов совмещения изображений путем максимизации взаимной информации не существует хороших способов оптимизации вычислений, как для корреляционных методов.

Итак, несмотря на все улучшения, методы, использующие низкоуровневые представления, не предоставляют универсального средства решения проблемы сопоставления изображений. Приемлемые решения имеются только для преобразований группы подобия, а уже для аффинных преобразований методы оказываются весьма ресурсоемкими и при этом неинвариантными по отношению к сложным преобразованиям яркости. В связи с этим низкоуровневые методы сопоставления имеют ограниченную сферу применения, и оказывается необходимым при решении задачи сопоставления изображений привлекать более высокоуровневые представления.