

第四章

人群移動行為的演化機制

在人群移動行為模擬的過程中，人群行走的路徑軌跡會因所設定的情境、地形與障礙物的變化而有所不同。例如在公園等戶外空曠場所，人群中個體間所保持的相對距離，相對於狹隘的都市街道來得更大，群體移動的速度也較為緩慢。由於隨其不同情境需求(例如空曠地方、單一出入口、活動空間中障礙物多的情境等等)，及不同的人群移動行為(例如群聚密集移動行為、疏散移動行為、守衛移動行為等等)，使得參數組合的最佳解也有所不同。一般而言，窮舉法無法有效找出此最佳解。因此本研究採用基因演算法的演化機制，自動找尋最佳解。基因演算法是一種模仿自然界天擇與優勝劣敗法則的搜尋方法。在演算的過程中，先將問題相關之參數進行編碼並計算各種解的目標函數值。透過編碼過的解進行交配(crossover)、突變(mutation)等演化機制，再根據各解的適應值的大小來選擇(selection)何者會繼續演化，何者將被消滅。此流程一代一代的進行，最後解答將朝向最佳化的結果發展。基因演算法利用整個解空間的群體解來搜尋，而非如傳統演算法由單一起始點搜尋，可以避免落入局部最佳解的情形。由於以上的特性，使得基因演算法不同於以往的搜尋方式，較容易在問題空間找到整體的近似最佳解。

因此我們將會在 4.1 節先介紹本篇對於找尋最佳解所應用的基因演算法，在 4.2 節中進一步來討論如何將第三章中所設計人群模擬系統與基因演算法兩者之間的流程，最後我們在 4.3 節中則會說明不同的移動行為所需要的適應函數設計。

4.1 基因演算法

基因演算法的基本理論是由 Holland 於 1975 年首先提出[16]，其基本精神在於仿效生物界中「物競天擇，適者生存」的自然演化法則。基因演算法的作法是將問題參數編碼為染色體(Chromosome)，代表解空間中的一點，透過多組的染色體在解空間中搜尋最佳解。演化過程中，利用疊代的方式執行選擇(Selection)、交配(Crossover)及突變(Mutation)運算來交換族群中染色體的資訊，藉此能夠避免在最佳化過程中陷入區域最佳解中及相對地加強搜尋全域最佳解的能力。大部份傳統的搜尋方式，基於某些決策法則只能考慮一個點，這些方法對於多峰值模式的搜尋空間而言，有收斂到局部最佳值的危險。但是基因演算法天生就是並行的，因為它可以在參數空間同時評估許多點，因此基因演算法可以降低收斂到局部最佳值的機會，並且更容易達到全域最佳值。另外，基因演算法僅關心每個參數集合所產生解的品質即可工作，不必像許多對於問題的結構及參數需要完整的知識。由於基因演算法不需要問題具體的資訊，因此比大部份的搜尋方式更有彈性。綜合言之，基因演算法與傳統的搜尋演算法之間有下列四點相異之處[9]

- 基因演算法中，參與運算的是問題解集合的編碼，而非問題的解；
- 基因演算法同時考慮一群可能解，而非單一解；
- 基因演算法採用適應函數(Fitness Function)資訊，而非導數(Derivative)或其他輔助知識。因此，基因演算法對於連續(Continuous)或非連續(Non-continuous)變數的系統皆可適用，不像傳統搜尋方法只能用連續變數的系統。
- 基因演算法利用機率式的轉換規則，而非明確的決定規則來引導搜尋方向。

以下我們分別介紹本文所採用的演化計算基本操作定義，以及說明演化計算的流程。

4.1.1 演算法流程說明

進行演化機制之前必須事先定義一些參數值，例如每個決策變數編碼成二進位字串的長度(Length of Binary Strings)、每個世代染色體產生的數量、交配率(Crossover Rate)、突變率(Mutation Rate)及疊代停止的限制條件。然而交配率和突變率是被當作決定基因運算子是否對一組染色體改變的門檻。一般而言，交配率及突變率都定義在介於0到1之間的實數值；另外，疊代停止限制條件可定義為疊代的次數上限。

每個染色體在進化過程中，是藉由適應值來量化其在環境中的競爭能力。染色體再生便是根據適應值(Fitness)來作為染色體的選取機率。本文採用『輪盤式選擇法』(Roulette Wheel Selection Method) [16]，模仿自然界的“適者生存”原則。利用適應值愈佳者在輪盤上所佔面積就愈大，被選取的機率就愈高，即可選擇較優良的染色體以演化下一代的族群，如此代代相傳演化下去，必然可以找到適應函數的近似最佳解。其演算法進行方式可分為以下4個步驟：(1)族群初始化、(2)評估模擬表現、(3)選擇表現良好的染色體，(4)進行基因演化操作(複製、交配、突變)。

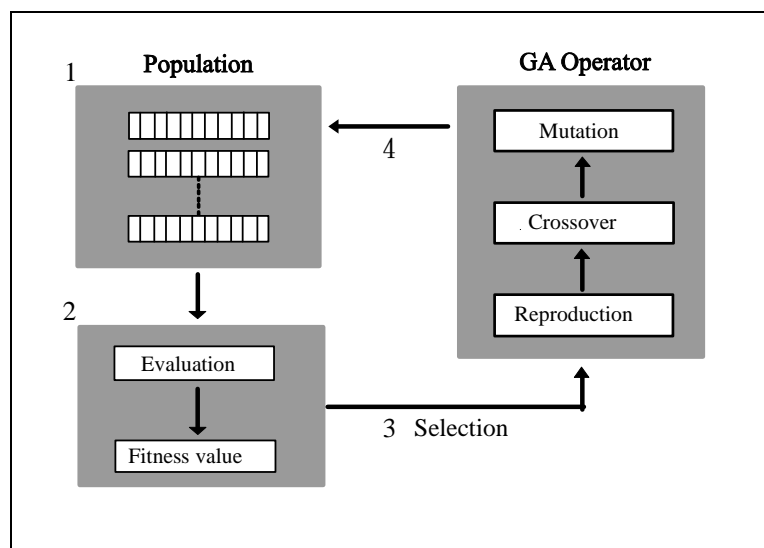


圖 4.1 基因演算法流程

4.1.2 編碼方式

基因演算法是將決策變數透過某種特定的方式，編碼成二進位字串，以『1』和『0』的組合表現狀態空間中某一個體的特徵(即染色體)，並以此進行運算。此一字串如同生物的染色體結構，字串本身的編碼長度即代表搜尋時解析度的高低；因此，在固定範圍的變數多維度空間中，搜尋間隔密度和染色體字串長度成正比。故在進行編碼過程時，就應慎選決策變數，並儘量減少編碼長度，以免浪費系統記憶空間，或產生系統溢載的情形。我們首先要確立每個參數的搜尋範圍，以便有效地搜尋整個狀態空間，再把每個參數以固定長度的字串加以編碼。染色體的編碼方式與其對應的適應評估函數為基因演算法在最佳化過程中最重要的關鍵。

4.1.3 適應函數

適應函數主要的目的是對一組經過編碼的可能解加以評分，而評分的依據是根據所處理的問題所訂定的。我們透過紀錄人群移動模擬的過程，作為適應函數所評估的對象。表現好的染色體則適應值越高，在之後選擇染色體的機率也越為增加，關於演化出不同的人群移動行為所需要的適應函數，也會隨不同移動行為有不同的適應函數組合，因此人群移動行為的演化可視為多目標最佳化的問題。我們將會在4.3節中詳述針對人群移動所設計的各项適應函數。

4.1.4 族群初始化

原始染色體族群的產生是基因演算法的第一步驟。通常越複雜的問題需要越大規模的族群來解決。基因演算法的特色之一是採用大量的可能解在解空間中搜尋，這許多經過編碼的可能解稱之為染色體族群。為了達到搜尋到全域最佳解的目的，族群初始化的方式

通常是以亂數產生可能解，使族群能夠均勻的散佈在整個解空間中。染色體族群最主要的目的是為了提供表現不同的個體。由於個體間的差異性，天擇的結果與基因交換的機制，才能產生下一代更好的染色體族群。如此週而復始，演化便持續的進行。產生初始族群之後，即進入演化的過程，其過程包含三個基本演化運算(Operators):複製、交配及突變。對於上述基本演化運算中的複製、交配及突變茲在4.1.5節分別詳述其動作。

4.1.5 基因演算法的基本運算元說明

基因演算法中的選擇與複製的機制，便是為了要模擬適者生存的運作模式。每個個體根據適應值的高低，決定該個體被再生的機率。適應值高的染色體，即有較高的機率被選中，進而自我複製出下一代的新染色體，此動作類似無性生殖。但是，適應值較差的染色體被選中而再生的機率，就會比適應值高的染色體還要小；因此新世代族群中，適應值差的染色體數量會比前一代族群更少，取而代之是適應值較佳的染色體。如此經過時代的演化，便能確保整個族群朝向最佳解的方向演化。基因演算法使用大量的可能解在解空間中平行搜尋，而其最大的優勢是能夠在平行搜尋的過程中彼此交換的資訊，而非漫無目標的搜尋，這就是交配運算的主要功能。

在染色體交配的過程中，透過交換彼此的部分資訊，期望能產生出新一代的染色體的適應度更加優秀。最常見的交配機制為單點交配(one-point crossover)。隨機決定一交配點(cut point)，將交配點後方的子字串(substring)交換，並得到兩個新的染色體。如果新的染色體具有較高的適應度，便可以將它置入原本的染色體族群之中。除了單點交配外，還有許多種交配的方式，如雙點交配(two-point crossover)、多點交配(multi-point crossover)、均勻交配(uniform crossover)等等。當然根據應用領域的不同，選擇適當的交配運算也是重要的課題之一。

根據上述的各種運算，基因演算法可以朝向近似最佳解的目標前進，但是若沒有其他的機制，基因演算法則會陷入區域最佳解中。因此，加上突變運算後，便可以確保在基因演算法演化的過程中，空間中的所有可能解都有被搜尋到的機會。突變運算是將交配後所繁衍的子代，根據預設的突變機率進行突變。突變的作法是把隨機選定的基因位元，由0轉1或1轉0。一般而言，突變機率值不能設太大，其主要目的是可以引進新的基因，使其在參數空間中隨意亂走，藉以開發新的搜尋區域，防止收斂至局部最佳點，可以探尋至全域最佳點。

4.2 人群移動的演化機制

在本研究中人群移動的演化機制，是利用 4.1 節中所討論的基因演算法。在基因編碼的部分是利用字串編碼方式(Bit String) 的方式，將第 3.2 章節所述驅動力的加權值參數組以二進位數字來表示，組成一條染色體。我們根據每個驅動力的加權值設計為基因(Gene)，由於虛擬力權重彼此相互影響的關係目前仍有許多討論空間，因此我們目前將虛擬力權重的解析度設定在小數點三位，分別針對各項虛擬力的權重加權值給予 10bits 的編碼，其染色體編碼後長度是 50 bits。編碼順序可參考 4.2 染色體編碼圖示。

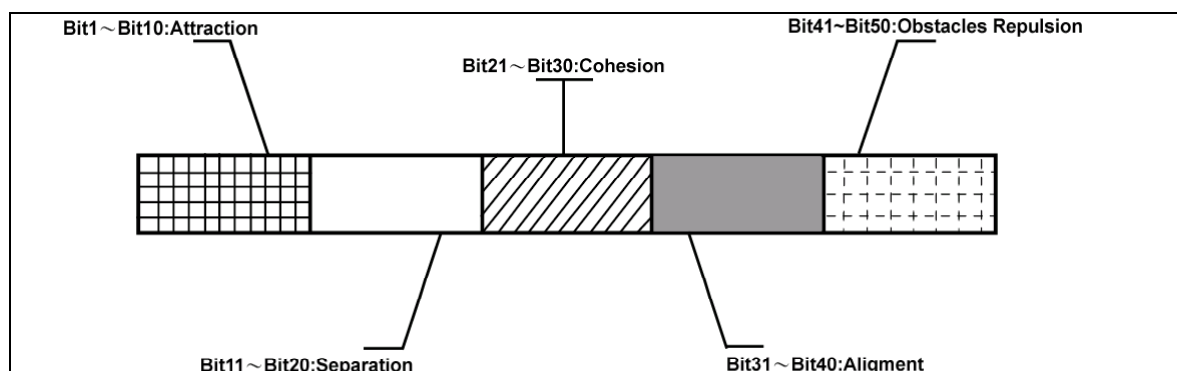


圖 4.2 染色體編碼

決定好編碼之後，便開始進行演化。我們演化人群所設定的代數為30代，每一代中產生100個染色體。每一個染色體套用於人群移動模擬中，對於虛擬力的權重設定，其每次模擬過程由系統紀錄，經適應函數的評估產生適應值(Fitness)。本研究的選擇機制是採用輪盤式選擇(roulette wheel selection)，在每一代的演化過程中，首先依每個物種(字串)的適應函數值的大小來分割輪盤的位置；適應函數值越大的話，則在輪盤上佔有的面積比例也越大。每個物種在輪盤上佔有的面積比例越大，代表被挑選到交配池中的機率越大。然後我們隨機選取輪盤的一點，將其所對應的物種選入到交配池中。本研究採用的交配機制為單點交配(one-point crossover)。隨機決定一交配點(cut point)。交換點是以亂數選擇介於1至個體長度之間的任意整數。將舊世代字串位於交換點右側的基因互相交換，生成兩個新染色體，完成交配運算。我們所採用的突變方法是隨機的選取一染色體的字串，並隨機選取突變點，改變染色體字串裡的位元資訊。突變的發生機率則由突變機率所控制。突變過程針對單一位元進行突變演算，對於二進制的位元字串就是將字串的0變1，1變成0。目前系統的突變機率值設定為1.5%，交配率為80%，此突變與交配的執行機率值的設定是參考先前研究以大量實驗所提出最適合的演化參數實驗結果[15]。

4.3 人群移動的適應函數定義

基因演算法的基本精神是以「適者生存」的模式，來搜尋最佳的問題解。使用者依其情景與移動行為的不同，所選取的適應函數也有所不同。這使得對於不同移動行為模擬所採用的適應函數為多個不同基本適應函數的組合。因此必須藉由正規化的動作，將多個不同的適應函數所分別得出的適應函數值，變成同一範圍內的適應值。配合使用者對於

各適應函數的加權比值高低的設定，並透過加權比重的總合產生適應值，以做為輪盤式選擇法決定染色體被選中而繁衍下一代的機率，以便搜尋出最佳解。

在討論適應函數設計前，我們首先必須定義虛擬人與人群的相對關係，本研究是採用以單一虛擬人為中心點，評估週遭鄰近虛擬人群與自身的相對關係。在圖 4.3 中，我們以 ID1 的虛擬人作為參考中心，並採用 k-nearest 的取樣機制。取樣目的在於由於個體有視野範圍的限制，因此便以視野範圍內的鄰近者視為可見鄰近者。我們參考 Miller [25] 在心理實驗中所得出的發現，人的記憶長度有 7 ± 2 的記憶單位，因此我們根據人類認知資源的有限性來設定取樣範圍。我們將系統 k 值取樣的設定為 7，在 ID1 視野範圍內的相對距離最近的週遭人群，其取樣參考人數最多可選 7 人。圖 4.3 中的淺灰色圈為視野範圍內的可見鄰近者群，在視野範圍內取樣出的 k 個鄰近者在圖中以較斜線淺色圈代表。在透過演化所產生出的模擬，系統需要針對此過程給予適應函數評估，以求得此次虛擬力參數組合基的表現好壞，作為之後演化的根據。在目前系統中，我們對於模擬的取樣是以人群的行進過程為主。因此，在人群到達最後系統所指定的最終範圍區時，之後的表現並不會列入評估。以圖 4.4 為例，圖中的軌道可視為模擬的取樣路徑(L)，在此路徑中的表現才會列入適應函數的評估。系統目前所定義的取樣範圍，主要可以避免在演化過程中，在人群移動到最後目的區域的時候，其模擬表現可以被排除計算，才不會發生所評估的區間都落在於最終人群移動到的目的區域，而稀釋了評分的效果。

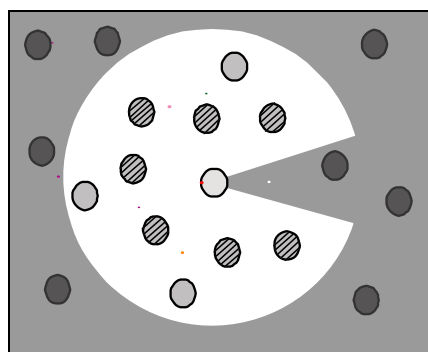


圖 4.3 鄰近群取樣圖

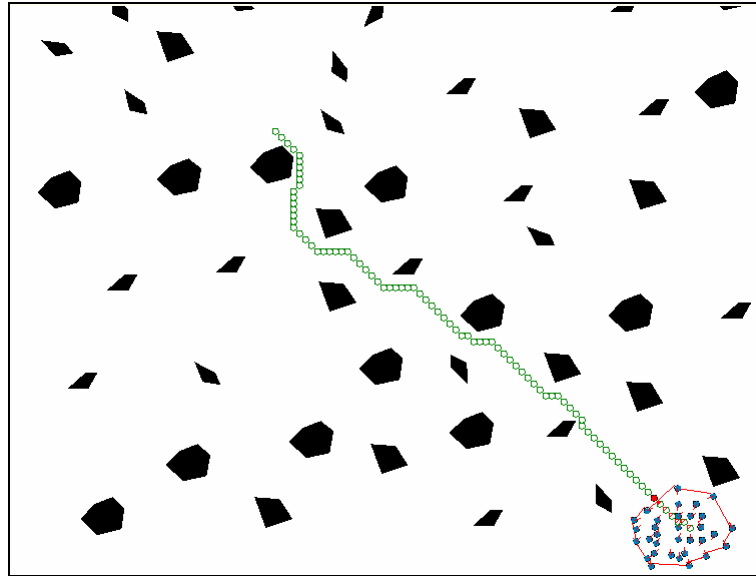


圖 4.4 模擬取樣範圍

在定義了個體如何取樣與其他鄰近者群的關係之後，以下分別針對本研究所定義的 5 項適應函數來說明。

4.3.1 群眾之間所保持的相對距離

如圖 4.5 所示，以個體(ID1)為中心找出k (在本例中k=7)個最近個體的相對距離($r_1, r_2, r_3, r_4, r_5, r_6, r_7$)，然後求其平均數(R_j) 如方程式(4.1)。與使用者所設定的期待值(U_e)相減，取其差額的絕對值。當差額越小時，表示其模擬的表現越能接近使用者的主觀期待，以(4.2)式所得出的值(G_j)，我們定義了 Q_d 為了做為正規化用途，使得 F_j 表線現在[0,1]之間，除以人群的個體數目(N)，以求平均值(R_m) (如(4.3)式)。假設總模擬時間長度為L，我們可求出此次模擬針對相對距離的適應函數值(G_a) (如(4.4)式)。

$$R_j = \frac{\sum_{i=1}^{i=k} r_i}{k} \quad (4.1)$$

$$G_j = \frac{Q_d}{|R_j - U_e|}, \quad 0 < |G_j| \leq 1 \quad (4.2)$$

$$R_m = \frac{\sum_{j=1}^{j=N} G_j}{N} \quad (4.3)$$

$$G_a = \frac{\sum_{m=1}^{m=N} R_m}{L} \quad (4.4)$$

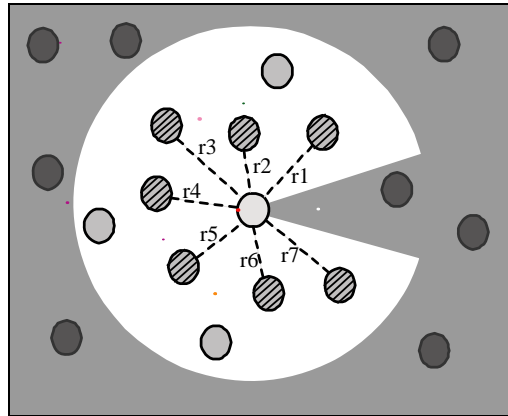


圖 4.5 人群間的相對距離

4.3.2 群眾與目標物的相對距離

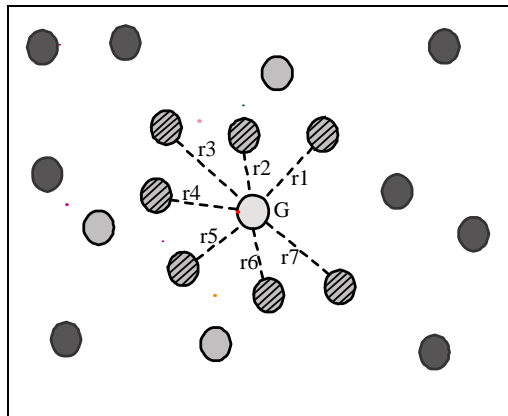


圖 4.6 群眾與目標物的相對距離

個體在某時間以目標終點為中心對於其所屬的人群求出兩者之間的相對距離 (r_i)，並以 k-nearest 鄰居的機制，在此群體中取樣出 k 個距離此目標物最近的鄰近者群，再求其平均數 (G_r) (如(4.5)式)，此為目標物與其他個體的相對距離 (G_m)。與使用者所設定的期待值 (U_e) 相減取絕對值，當差額越小時，表示其模擬的表現越能接近使用者的主觀期待 (G_j)

(如(4.6)式所示) 同樣我們為了做為正規化定義了 Q_e ，使得 G_j 表線現在 $[0,1]$ 之間，。群體的模擬表現是以人群的個體數目(N)人求平均值(G_k) (如(4.7)式)。假設總模擬路徑時間長度為 L ，則在此次模擬針對彼此相對距離的適應函數值(G_B)可如(4.8)式所定義。(請參考圖 4.6 範例所示。)

$$G_m = \frac{\sum_{i=1}^{i=k} r_i}{k} \quad (4.5)$$

$$G_j = \frac{Q_e}{|G_m - U_e|}, \quad 0 < |G_j| \leq 1 \quad (4.6)$$

$$G_k = \frac{\sum_{j=1}^{j=N} G_j}{N} \quad (4.7)$$

$$G_B = \frac{\sum_{k=1}^{k=L} G_k}{L} \quad (4.8)$$

4.3.3 模擬過程中人群移動所發生的碰撞次數

記錄此群體在每次模擬路徑中的碰撞次數。碰撞次數越低，代表表現越好。個體在單位時間中記錄此團體中虛擬人群所發生的碰撞次數(C_s)。我們將團體人數(N)減去發生的碰撞次數之後，經由正規化後所得即為適應函數值 C_j (如(4.9)式)。系統再依照總模擬路徑的長度為 L ，計算得出此人群的碰撞適應函數值 G_D (如(4.10)式)。

$$C_j = \frac{N - C_s}{N} \quad (4.9)$$

$$G_D = \frac{\sum_{j=1}^{j=N} C_j}{L} \quad (4.10)$$

4.3.4 人群面向角度的一致性

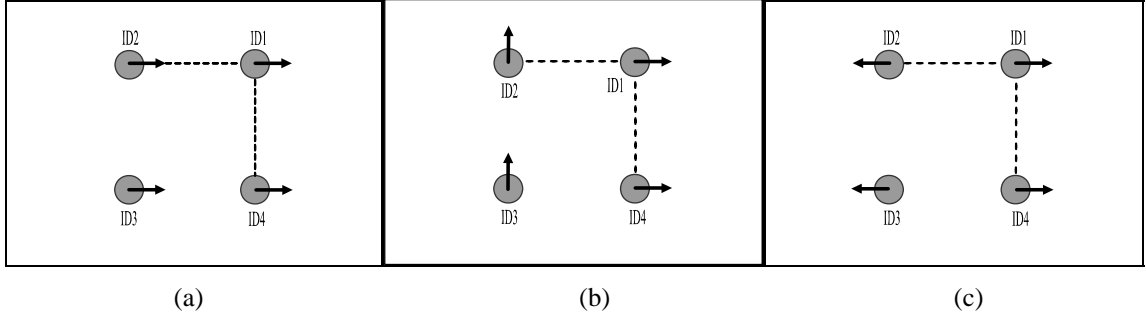


圖 4.7 群體角度的一致性

群體移動方向的一致性，在我們的研究提出以個體與周遭鄰近群面向的差異來討論。因此在模擬過程中，我們以自身面向角度與週遭人群面向角度的差異來定義此適應函數。此差異量越小表示其面向角度的一致性越高。以在圖 4.7(a)中的ID1 為例，透過k-nearest 鄰居的方式取樣出ID2 及ID4，再計算ID1 與ID2 及ID4 之間面向角度的差異量(如(4.11)式所示)，代入(4.12)式即可得出個體一致性的適應函數值。在圖 4.6 的範例中，一致性次高者為圖 4.7(b)，而一致性最低者為圖 4.7(c)。對於群體的一致性表現，便是根據上述做法求出每一個體的局部一致性，再求其在單位時間內的群體所得到平均的一致性表現(如(4.13)式所示)。最後再根據模擬的時間長度求出整體對於一致性表現的適應函數值 G_c (如(4.14)式所示)。

$$A_j = \frac{\sum_{i=1}^{i=k} (C_j - C_i)}{k}, \quad i \neq j \quad (4.11)$$

$$A_j = \left(1 - \frac{\theta_j}{180}\right) \quad (4.12)$$

$$A_m = \frac{\sum_{j=1}^{j=N} A_j}{N} \quad (4.13)$$

$$G_c = \frac{\sum_{m=1}^{m=L} A_m}{L} \quad (4.14)$$

4.3.5 人群相對距離的恆常性

關於人群相對距離的恆常性，我們是以虛擬人與群體彼此之間距離的變異量為主要考量。如圖 4.8(a)所示，以個體(ID8)為中心，取樣出k個距離此個體最近的個體(ID4, ID7, ID9, ID10)，再求其個體與其他個體的相對距離(r_i)，並求出平均數(R_{mean}) (如(4.15)式所示)。以ID8 分別對於鄰近個體求出的相對距離(r_i)與平均數(R_{mean})，再以此算出相對距離的變異量(如(4.16)式所示)。代入(4.17)式可得到ID8 與週遭鄰近群的相對距離變異量的適應函數值(H_j)，同樣我們為了做為正規化定義了 Q_f ，使得 F_j 表線現在[0,1]之間。群體在單位時間的模擬表現是以人群的個體數目(N) 求出群體的平均適應函數值(F_i) (如(4.18)式所示)。配合總模擬路徑時間長度 L ，可以利用(4.19)式算出群體對彼此相對距離變異量的整體適應函數值(G_E)。從圖 4.8(b)中，對ID1,2,3,4,6,8,10,12 而言，距離變異量為 0，但對ID5,7,9,11 而言，其不為零的變異量將增加群體的變異量，並減少了群體相對距離一致性的適應函數值。

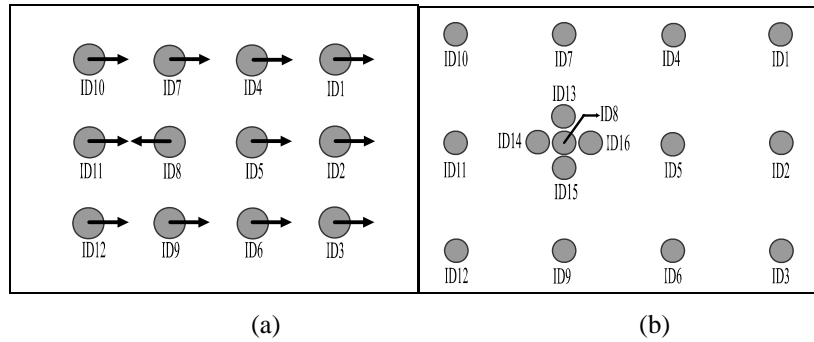


圖 4.8：相對距離的恆常性

$$R_{mean} = \frac{\sum_{i=1}^{i=k} R_i}{k} \quad (4.15)$$

$$S_j = \frac{\sqrt{\sum_{i=1}^{i=k} (R_i - R_{mean})^2}}{k} \quad (4.16)$$

$$H_j = \frac{Q_f}{|S_j|}, \quad 0 < |H_j| \leq 1 \quad (4.17)$$

$$V_i = \frac{\sum_{j=1}^{j=N} H_j}{N} \quad (4.18)$$

$$G_E = \frac{\sum_{t=1}^{t=L} V_t}{L} \quad (4.19)$$

透過上述適應函數所產生的適應值，由於每項適應函數的設計皆為不同，因此我們分別根據適應函數所表現的合理範圍進行正規化方式，並且取得正規化之後的各項適應值。本研究中所採用的正規化方式為按照各項本身適應函數的變化，取其合理的上下限區間範圍，以轉換為相同的尺度，使得適應值之間可以相互比較，並根據不同移動行為來採取不同的適應函數權重組合。本系統以不同的加權指數 (W_1, W_2, W_3, W_4, W_5)，計算一個行為的總適應值(F_{sum})如方程式(4.20)，以做為是模擬人群移動所產生的演化適應值。至於如何針對不同的移動行為來制訂所屬的加權指數，我們是透過設計者主觀的設計，針對使用者對於特定行為來設定適應函數的權重，比設定虛擬力權重來得更完直觀。這有別於上述所提及將虛擬力權重以演化方式來決定，主要這部分將會在第五章實驗設計來說明。

$$F_{sum} = W_1 * G_A + W_2 * G_B + W_3 * G_C + W_4 * G_D + W_5 * G_E \quad (4.20)$$

在本章節中我們定義了人群移動行為如何透過演化方式來進行，並且根據個體的特性來設計對應的適應函數，其中尤以對於取樣部分我們透過實驗方法決定了個體與周遭鄰近群的取樣，以及總體在表現過程中如何取出有效的模擬範圍以做為適應函數評估所用。我們將提出如何以演化的適應函數設計來評估各種行為的表現，並演化出特定行為所對應的虛擬力加權組合。在本研究中，我們以基因演化的方式，我們只需設定好初始人群狀態，針對個體區在域性表現所表現出群體移動過程的狀態，給予適應值分數的評估。因此透過演化方式便可演化出人群彼此區域性相互影響所產生出來的穩定結果，而不需要針對群體的狀況做全區域考量，可避免落入整體最佳化的求解複雜度。