

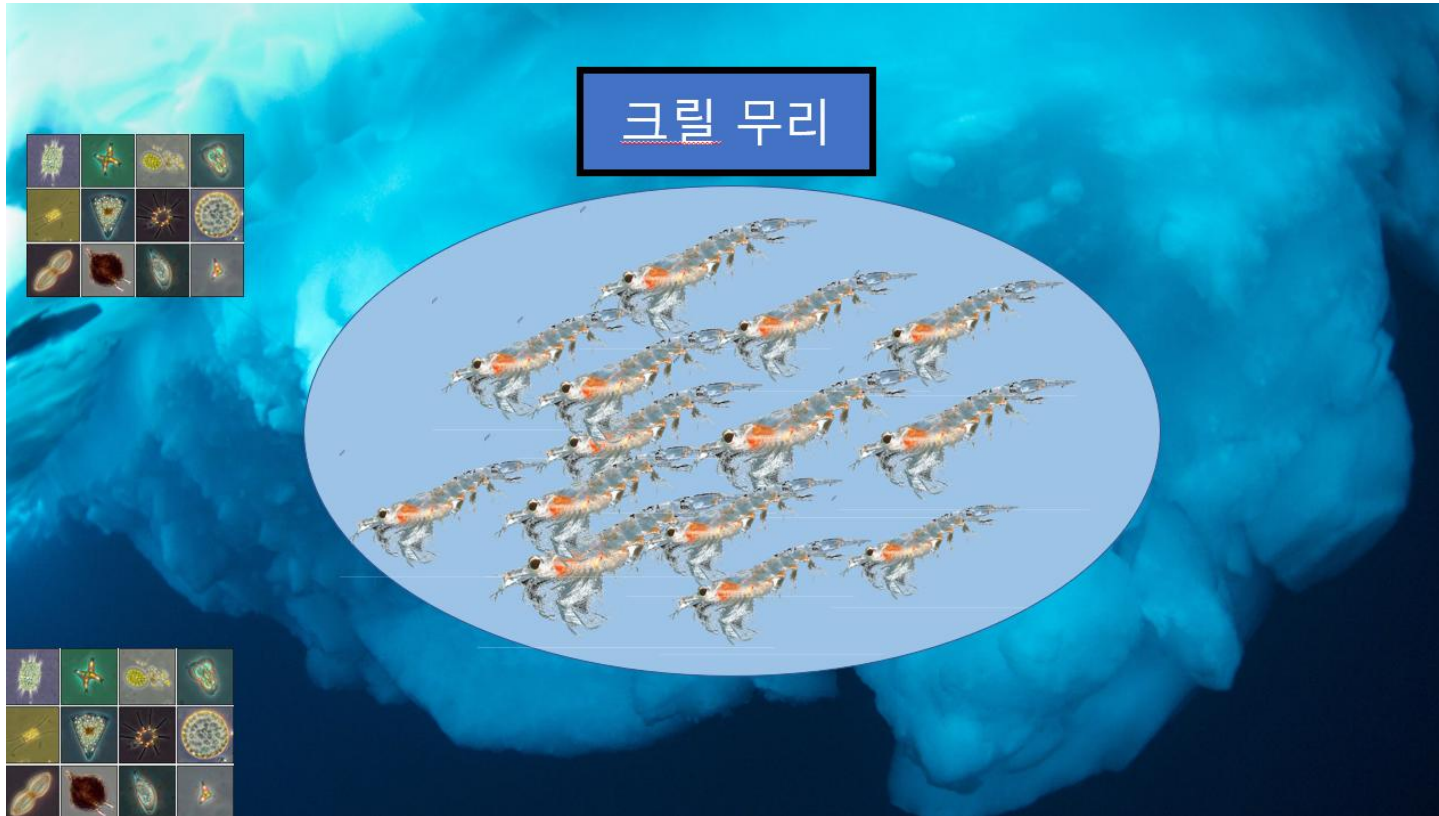
크릴허드 알고리즘을 이용한 피쳐 선택션

Feature Selection using Krill Herd algorithm

이후현

크릴허드 알고리즘이란?

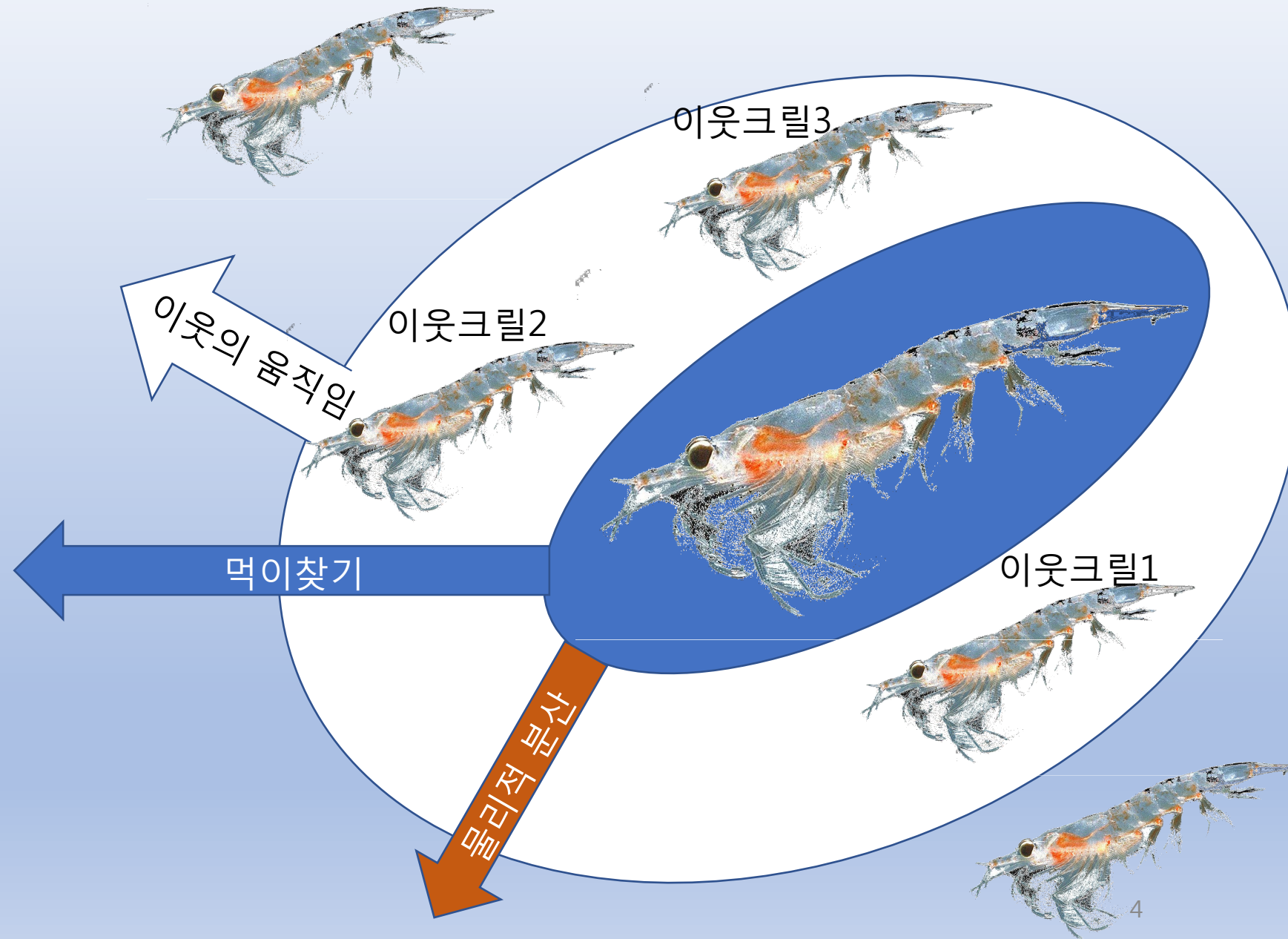
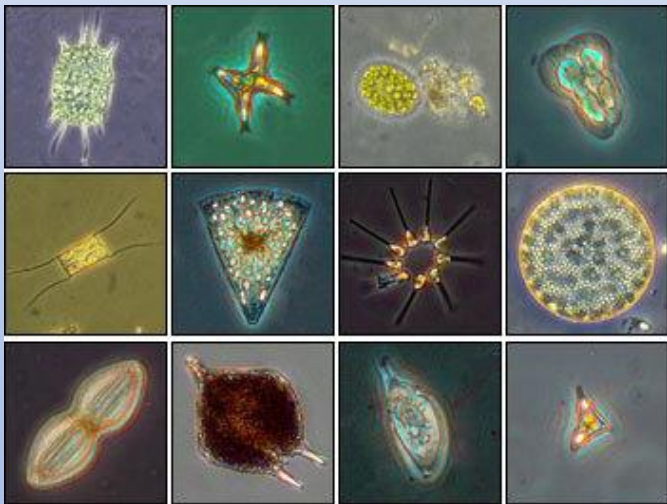
- 크릴 새우 무리의 움직임을 모방한 생태계 모방 알고리즘



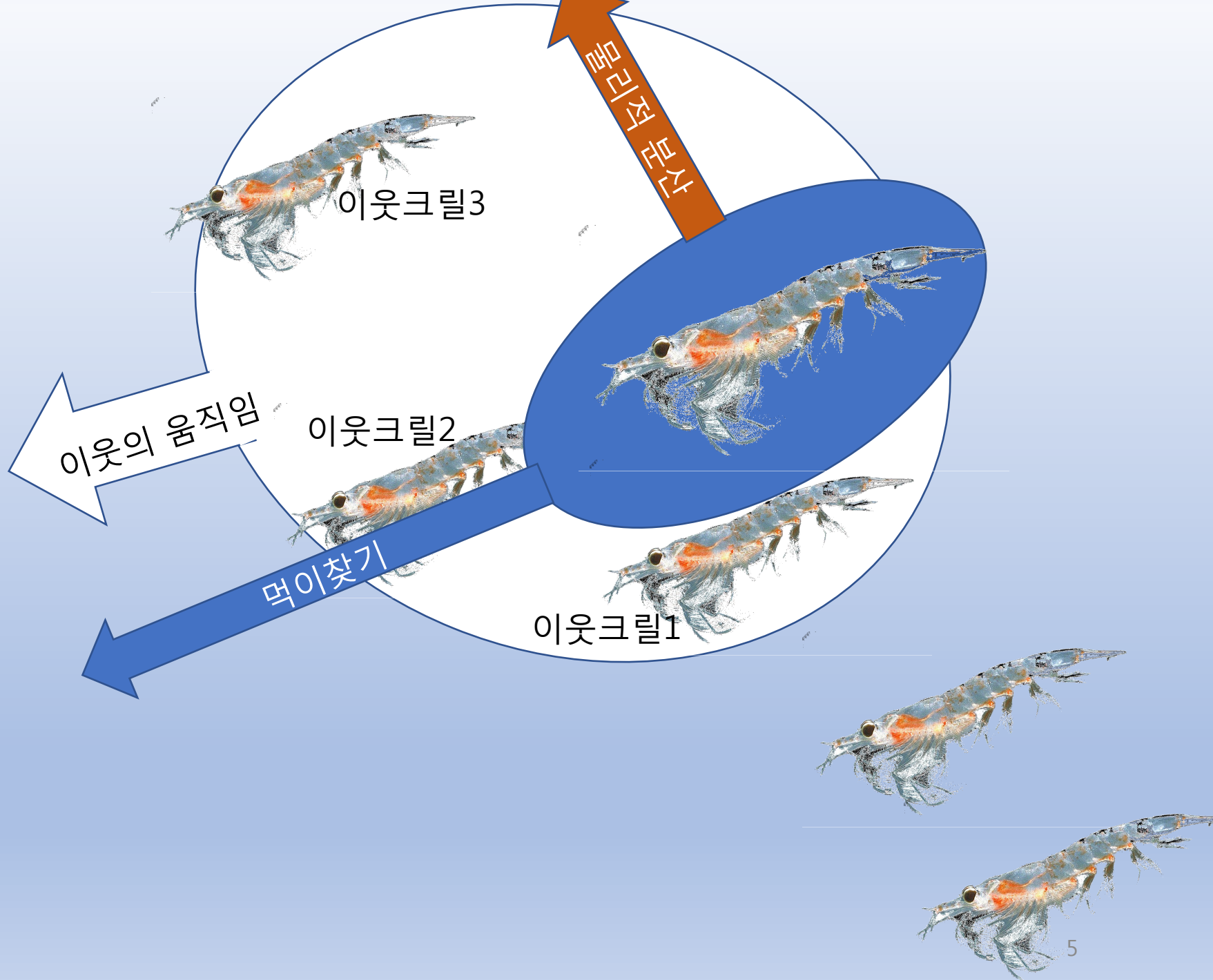
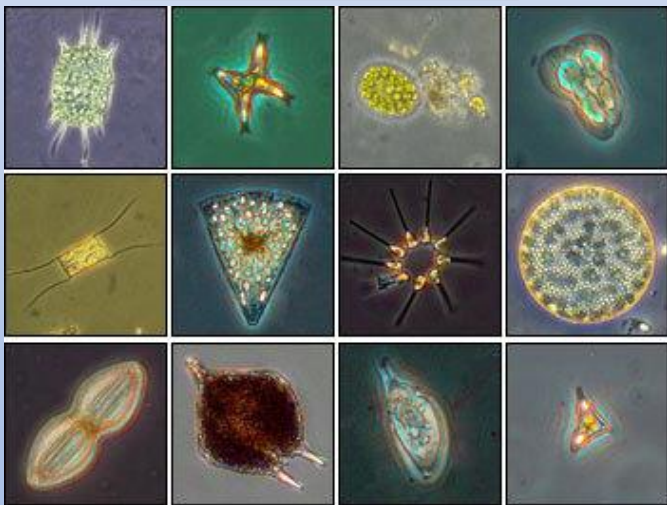
크릴허드 알고리즘의 동작

- 무리 안의 크릴 새우의 움직임을 결정하는 것들
 - 이웃 새우들의 움직임
 - 먹이의 위치
 - 각 개체의 임의의 물리적 분산

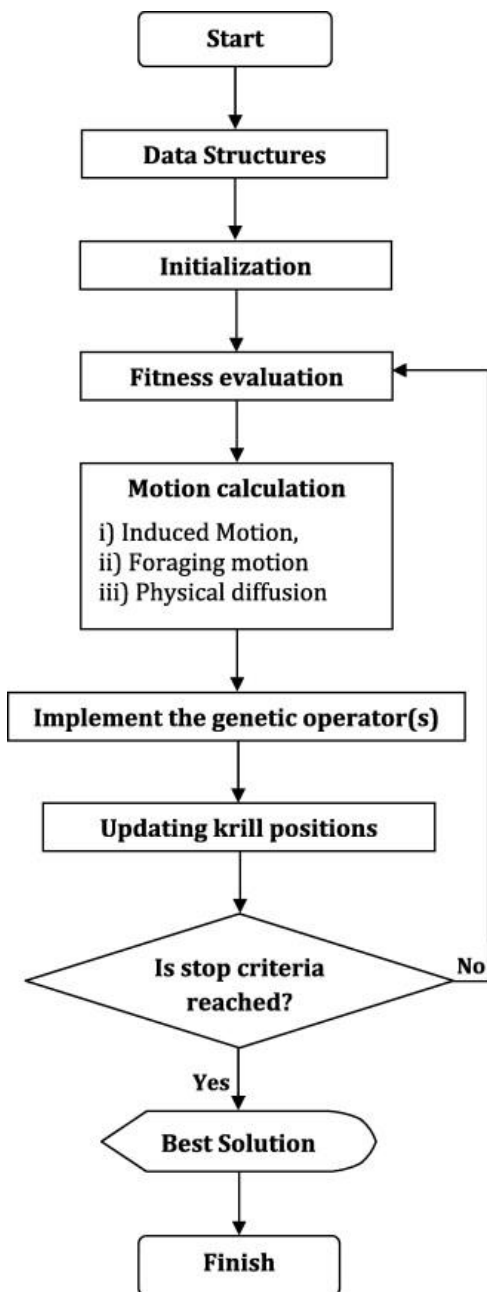
먹이



먹이



크릴허드 알고리즘의 동작



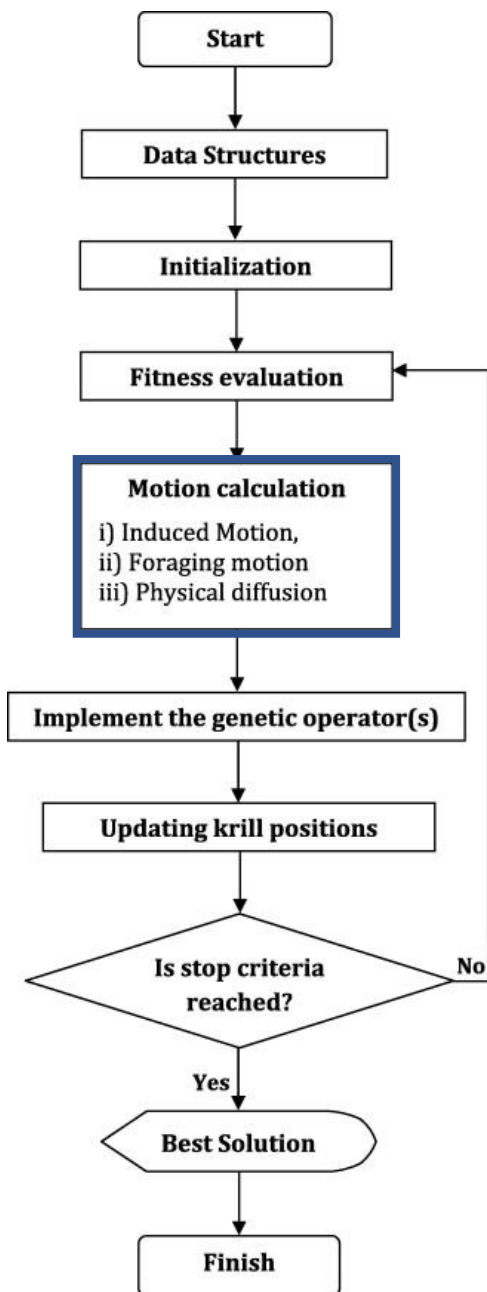
1. Data Structures : 알고리즘의 매개변수와 기타 동작에 필요한 값을 정의
2. Initialization : 크릴을 탐색공간 임의의 위치에 배치
3. Fitness evaluation : 크릴의 위치를 기반으로 적합성을 평가
4. Motion calculation : 크릴의 움직임을 계산
5. Implement the genetic operator(s) : 유전 알고리즘을 이용하여 섞음
6. Updating krill positions : 4번과 5번으로 계산한 새로운 위치로 크릴의 위치를 이동
7. 조건을 만족할 때 까지 3~6을 반복

크릴허드 알고리즘의 동작

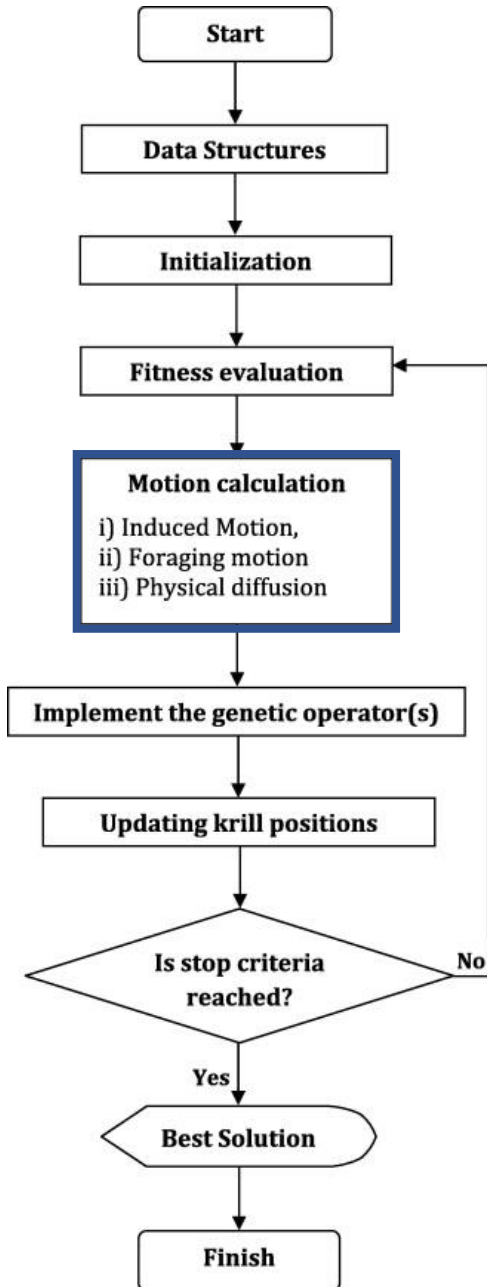
4. Motion calculation

$$\frac{dX_i}{dt} = N_i + F_i + D_i$$

- X_i : 크릴의 위치
- N_i : 유도되는 움직임(Induced motion)
- F_i : 수렵 활동(Foraging motion)
- D_i : 물리적 분산(Physical diffusion)
- i : 반복 횟수(iteration)



크릴허드 알고리즘의 동작



4. Motion calculation

- Induced motion

- 크릴 개체는 이웃 크릴과의 상호작용을 통해 높은 밀도를 유지하고 움직임

- $$N_i^{new} = N^{max} \alpha_i + w_n N_i^{old}$$

- $$\alpha_i = \alpha_i^{local} + \alpha_i^{target}$$

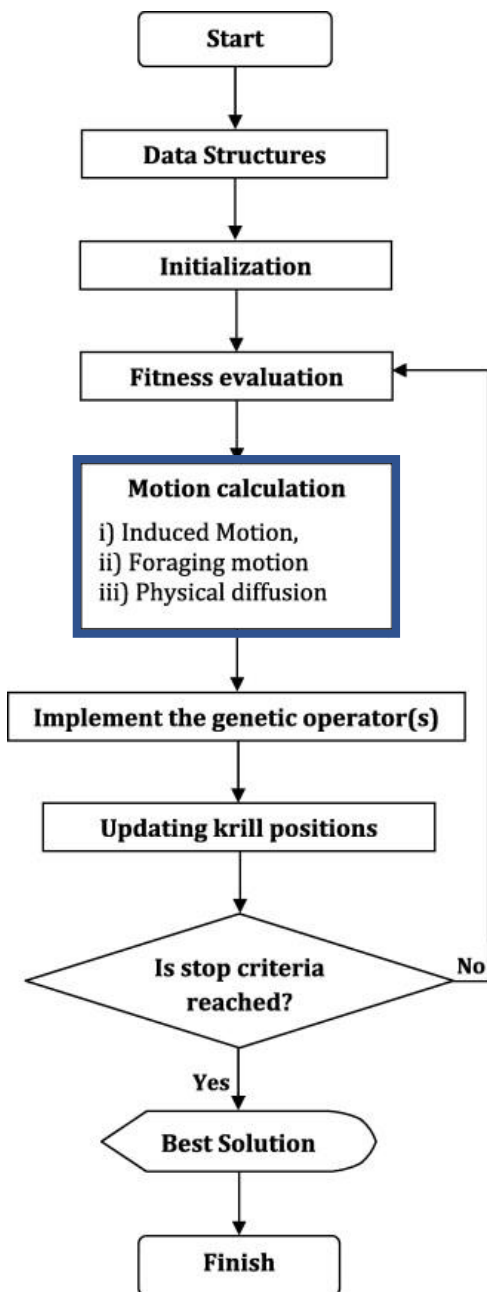
w_n : N_i^{old} 를 적절하게 반영하기 위한 무게중심[0,1]

N^{max} : 유도되는 움직임의 최대속도(0.01)

α_i^{local} : 가까운 이웃 크릴에게 받는 영향

α_i^{target} : 모든 크릴 중, 장 적합성이 높은 크릴의 위치로 향하는 움직임

크릴허드 알고리즘의 동작



4. Motion calculation

• Foraging motion

- 수렵 활동은 음식의 위치와 음식 위치에 대한 이전의 경험을 기반으로 수행

- $F_i = V_f \beta_i + w_f F_i^{old}$

- $\beta_i = \beta_i^{food} + \beta_i^{best}$

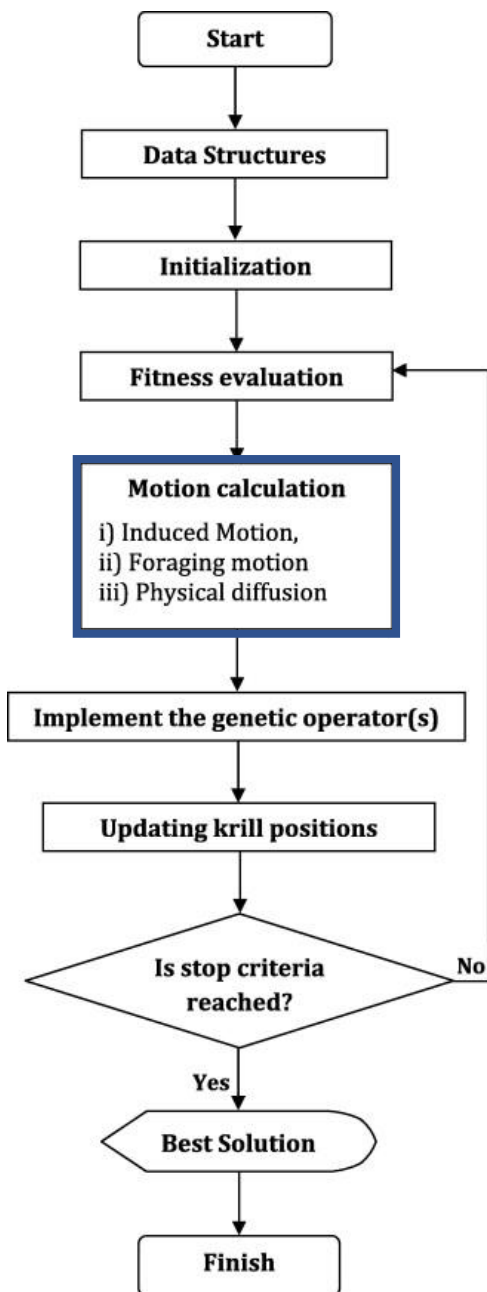
w_f : F_i^{old} 를 적절하게 반영하기 위한
무게중심[0,1]

V_f : 실제 크릴이 수렵활동 하는 속도(0.02)

β_i^{food} : 실제 음식이 있을 것으로 여겨지는 위치

β_i^{best} : 크릴이 가장 좋은 적합성을 보여주었던
개체 과거의 위치로 향하는 움직임

크릴허드 알고리즘의 동작



4. Motion calculation

- Physical diffusion

- 크릴 새우에게 무작위적인 움직임을 부여하는 과정

- $$D_i = D^{max} \left(1 - \frac{I}{I_{max}}\right) \delta$$

D^{max} : 최대 diffusion 속도[0.002,0.010]

δ : 임의의 방향벡터

I : 현재 Iteration

I_{max} : 최대 Iteration

피쳐 선택

feature1	feature2	feature3	feature4	Label
2	5	3	3	0
3	4	8	0	1
2	1	9	3	1

feature1	feature2	feature3	feature4
1	0	1	0

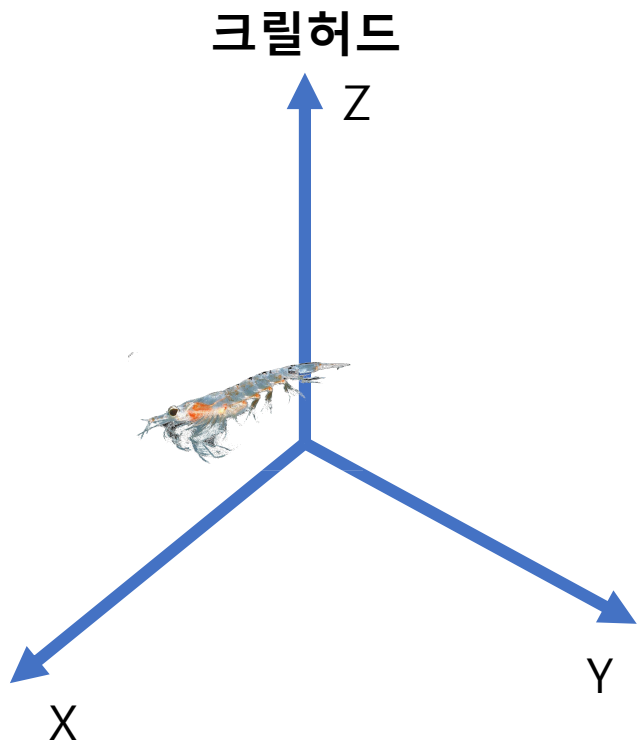
사용할 피쳐 선택

불필요한 feature를 제거

- 정확도 향상
- 계산시간 향상
- Feature를 얻기 위한 자원소모 감소

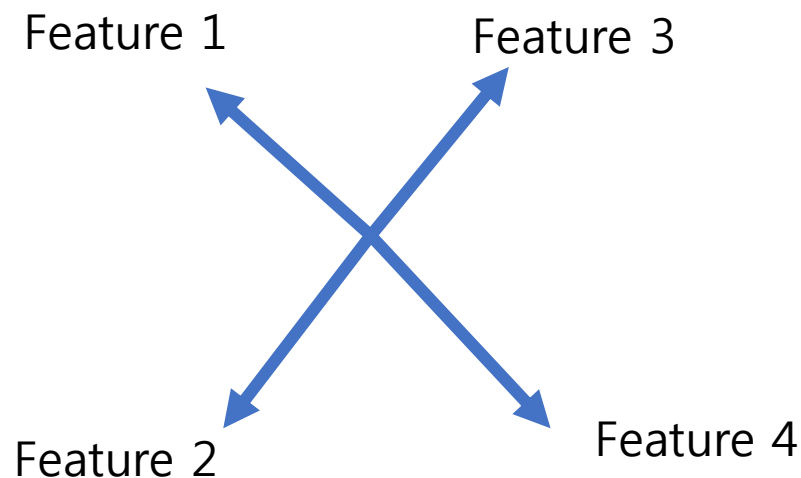
feature1	feature3	Label
2	3	0
3	8	1
2	9	1

크릴허드 알고리즘을 이용한 피쳐 선택션



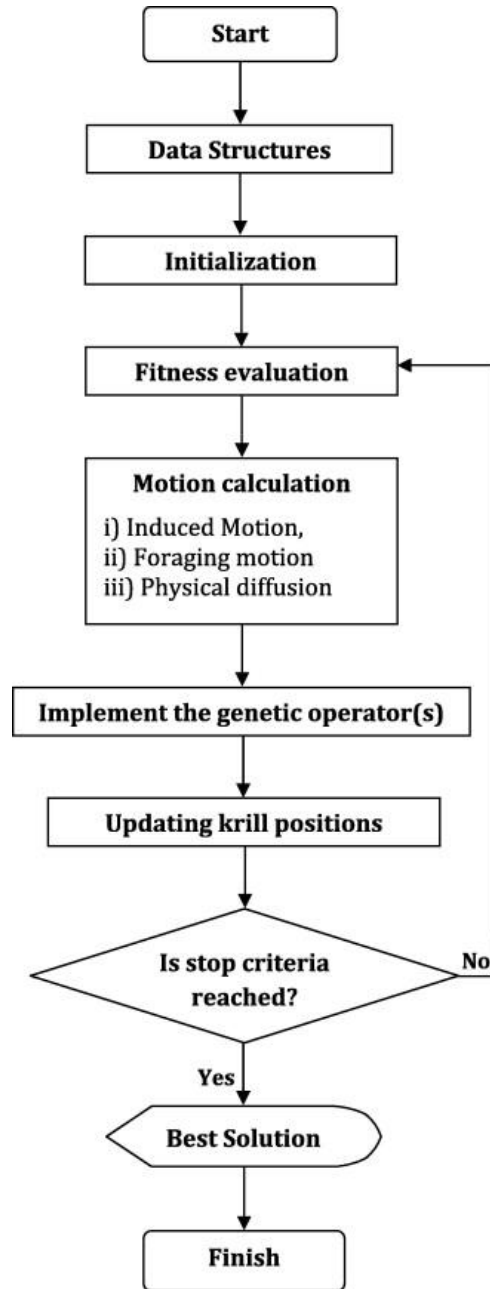
- 크릴이 움직이는 3차원 장소
- 연속된 위치 값을 가질 수 있다.
- 먹이

KH를 이용한 피쳐 선택션



- 원 데이터셋의 feature만큼의 차원
- Feature Selector는 0이나 1만을 갖는다
- 적합성(fitness)

피쳐 셀렉션 알고리즘의 동작



1. Data Structures : 알고리즘의 매개변수와 기타 동작에 필요한 값을 정의
ex) w_f , V_f , w_n , N^{max} 등
2. Initialization : 각 크릴을 임의로 생성한 피쳐 셀렉션 공간에 놓음
3. Fitness evaluation : 각 크릴의 피쳐 셀렉션으로 선택한 데이터로 학습시키고(msvm) 그 적합성 평가
4. Motion calculation : 크릴의 움직임을 계산
5. Updating krill positions : 4번과 5번으로 계산한 새로운 위치로 크릴의 위치를 옮김.
 - 계산된 위치는 0에서 1사이의 숫자
 - 그 숫자를 기반으로 확률적으로 0과 1을 선택
6. 조건을 만족할 때 까지 3~6을 반복

결과

	Selection Value	Selection Time	Normal Value	Normal Time	선택한 feature의 개수
BASEHOCK	0.9629	0.0216	0.9679	0.217	4172/4862
Carcinom	0.8076	0.062	0.8082	0.0978	7549/9182