

## 生物工学 Ⅱ

次の[Ⅱ - 1]～[Ⅱ - 3]の 3 題を、それぞれ別の解答用紙に答えよ。

## [Ⅱ - 1]

DNA シーケンシング技術の発達により、500～700 塩基程度の長さの DNA 断片の塩基配列をほぼ自動的に決定し、得られたデータから計算機を用いてゲノム全体を再構築できるようになった。得られた塩基配列は国際塩基配列データベースに登録され、誰でも制限なしに利用可能となっている。2012 年 6 月現在、国際塩基配列データベースに登録されている全てのエントリの合計塩基数は 1400 億を超えており、その中から特定の塩基配列を探索することは、計算機の利用なくしてはほぼ不可能である。

塩基配列は、数種類の文字種からなる文字列として表現することができるため、塩基配列の探索には、計算機による文字列探索技術が適用できる。計算機による文字列探索とは、テキストと呼ばれる長さ  $n$  の文字列と、パターンと呼ばれる長さ  $m$  の文字列が与えられたときに、テキスト中にパターンが現れるかどうかを判定し、現れる場合には、テキスト中でパターンが始まる位置を求める処理である。本問では、テキストはパターンよりも長く、 $n > m$  であるとする。

以下では、テキストおよびパターンの各文字を  $T_i$  および  $P_i$  として表す。長さ  $n$  のテキストは  $T_1 T_2 \dots T_n$ 、長さ  $m$  のパターンは  $P_1 P_2 \dots P_m$  と表される。テキストの  $i$  番目の文字からパターンを照合する場合の、テキストとパターンの対応関係を図 1 に示す。また文字種は断りのない限り A, G, C, T の 4 種類であるとする。

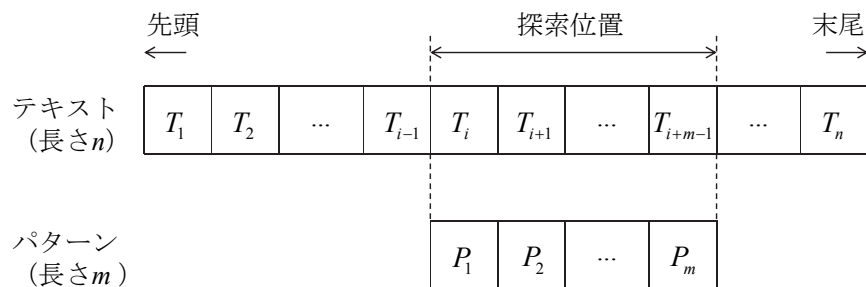


図 1

(問 1) 最も単純な文字列探索方法として、力まかせ法と呼ばれる手法がある。力まかせ法では、まずテキスト内にパターンの最初の文字  $P_1$  が出現する位置を、テキストを先頭から 1 文字ずつ  $P_1$  と照合して探索する。テキストの  $i$  番目の文字  $T_i$  が  $P_1$  と一致した場合、続いて  $T_{i+1}$  から  $T_{i+m-1}$  までの文字が、それぞれ  $P_2$  から  $P_m$  と一致するかを順次照合する。途中で不一致が生じた場合、以降の照合を中止し、文字  $P_1$  の探索をテキストの  $i+1$  番目の文字  $T_{i+1}$  から再開する。パターンの文字が対応するテキストの文字と全て一致した場合、パターンが発見できたことになる。

(次のページに続く)

- (ア) 力まかせ法を用いてテキスト GTCGATGTGCGTGACCG からパターン GTGCAC を発見するまでの文字照合過程を全て示し、文字照合回数と、そのうち文字が一致した回数を答えよ。
- (イ) 力まかせ法を用いて、長さ  $n=10$  のテキストから長さ  $m=5$  のパターンを探索する場合に、文字照合回数が最大となるテキストとパターンの組み合わせを一つ示せ。ただし、テキストの部分文字列  $T_6T_7T_8T_9T_{10}$  がパターン  $P_1P_2P_3P_4P_5$  と一致するものとする。テキストとパターンのいずれも、4 種類の文字種を全て含んでいる必要はない。

(問 2) 力まかせ法では、テキストのある位置からパターンを照合した結果が不一致であった場合に、探索位置を 1 文字だけ末尾の方に移動させる。これに対して、テキストのある位置からパターンを照合する際に、パターンの末尾の文字から照合を行うと、不一致が生じた場合に、パターンに関する知識に基づいて探索位置を一度に大きく移動させることができ、文字照合回数の削減が期待できる。この方法を本問では末尾探索法と呼ぶ。

例として、テキスト ATGAGTGATGTA から、パターン CCC を探索する事例 1 を考える。力まかせ法と同様に、探索位置はテキストの先頭から末尾に向かって移動させる。最初に  $T_1T_2T_3$  (ATG) と  $P_1P_2P_3$  (CCC) を照合するが、力まかせ法と異なり、まずパターンの末尾の文字  $P_3$  (C) を  $T_3$  (G) と照合する。 $T_3$  (G) と  $P_3$  (C) が異なることから、 $T_1T_2T_3$  (ATG) は  $P_1P_2P_3$  (CCC) と一致しないことが 1 文字の照合で確定する。加えて、 $T_3$  すなわち G はパターン  $P_1P_2P_3$  (CCC) 中に出現しないことに着目すると、 $T_3$  を含むテキストの部分文字列がパターンと一致する可能性はないことも判明する。よって、次は ( ① ) と  $P_1P_2P_3$  の照合まで探索位置を移動させることができる。

別の例として、事例 1 と同じテキストから、パターン GCC を探索する事例 2 を考える。事例 1 と同様に、 $T_3$  (G) と  $P_3$  (C) が異なることから  $T_1T_2T_3$  (ATG) は  $P_1P_2P_3$  (GCC) と一致しないことが確定する。事例 1 との違いは、 $T_3$  すなわち G がパターン  $P_1P_2P_3$  (GCC) に含まれているため、 $T_3$  とパターン内の G が対応する探索位置で、テキストの部分文字列とパターンが一致する可能性がある点である。したがって次は ( ② ) と  $P_1P_2P_3$  を照合する必要があることがわかる。

以上の事例を参考にすると、末尾探索法の戦略の一つとして、不一致が生じたテキストの文字に着目し、その文字に一致するようにパターンの移動量を決定する方法が考えられる。具体的には、テキストの部分文字列  $T_iT_{i+1}\dots T_{i+m-1}$  とパターン  $P_1P_2\dots P_m$  の照合において、不一致がパターンの先頭から  $j$  文字目の照合、すなわち  $P_j$  と  $T_{i+j-1}$  の照合で生じたとすると、パターンの不一致位置より先頭側の部分文字列  $P_1P_2\dots P_{j-1}$  において  $T_{i+j-1}$  と一致する最後の文字が、 $T_{i+j-1}$  と比較されるように探索位置を移動すればよい。つまり、 $T_{i+j-1} = P_k (1 \leq k \leq j-1)$  となる最大の  $k$  を発見し、探索位置を ( ③ ) 文字だけ末尾の方に移動させる。また、そのような  $k$  が存在しない場合には、探索位置を ( ④ ) 文字だけ末尾の方に移動させる。

- (ア) 末尾探索法により文字照合回数の削減が期待できることを、次のようなテキストとパターンを用いて確認してみる。この小問 (問 2 (ア)) においてのみ文字種を 26 種類とする。テキスト ABCDEFGHIJKLMNOPQRSTUVWXYZ にパターン ZZZZZZ が含まれていないことを判定するために照合されるテキストの文字を、照合された順に全て挙げよ。

(次のページに続く)

(イ) ( ① ) および ( ② ) に入るテキストの部分文字列を, 以下の(a)~(d)の中からそれぞれ一つ選択せよ.

(a)  $T_2T_3T_4$  (b)  $T_3T_4T_5$  (c)  $T_4T_5T_6$  (d)  $T_5T_6T_7$

(ウ) ( ③ ) および ( ④ ) に入る式を  $i, j, k$  および数字を用いて書け.

(エ) 末尾探索法を用いてテキスト GTCGATGTGCGTGACCG からパターン GTGCAC を発見するまでの文字照合過程を全て示し, 文字照合回数と, そのうち文字が一致した回数を答えよ.

(オ) 末尾探索法を用いて, 長さ  $n=10$  のテキストから長さ  $m=5$  のパターンを探索する場合に, 文字照合回数が最大となるテキストとパターンの組み合わせを一つ示せ. ただし, テキスト内にパターンは含まれていないものとする. テキストとパターンのいずれも, 4 種類の文字種を全て含んでいる必要はない.

(カ) 末尾探索法においてパターンに関する知識を利用するための別の戦略として, テキストとパターンの不一致が生じた場合に, それまで一致していた部分に関する情報を利用して, 探索位置の移動量を決定する方法も考えられる. テキストの部分文字列  $T_iT_{i+1}\dots T_{i+m-1}$  とパターン  $P_1P_2\dots P_m$  の照合において, 不一致がパターンの先頭から  $j$  文字目の照合, すなわち  $P_j$  と  $T_{i+j-1}$  の照合で生じたとして, この戦略の具体的なアルゴリズムを設計せよ.

[Ⅱ - 2]

図 1(a)(b) に示す剛体が  $xy$  平面内を  $x$  軸方向へ転がる運動を考える．図 1(a) の剛体は，厚さを無視できる半径  $r$  の円環であり，その質量を  $m$  とする．図 1(b) の剛体はリムレスホイールと呼ばれるもので，厚さを無視できる半径  $r_0$  の円環に，その中心から放射状に広がる長さ  $l_0$  の細いスポークが取り付けられている．ここでは，そのスポークを脚と呼び，隣り合う脚の成す角を  $2\alpha$  とする．脚の数を  $N$  とすると， $\alpha = \pi/N$  であり，図 1(b) は  $N = 10$  の例である．円環部分の質量を  $m$  とし，脚の太さおよび質量は無視できるとする．地面に接し，その接点が回転の軸になる脚を支持脚と呼び，その他の脚を遊脚と呼ぶ．なお，重力加速度を  $g$  とする．また，2つの円環の密度は一樣とする．このとき，以下の設問に答えよ．

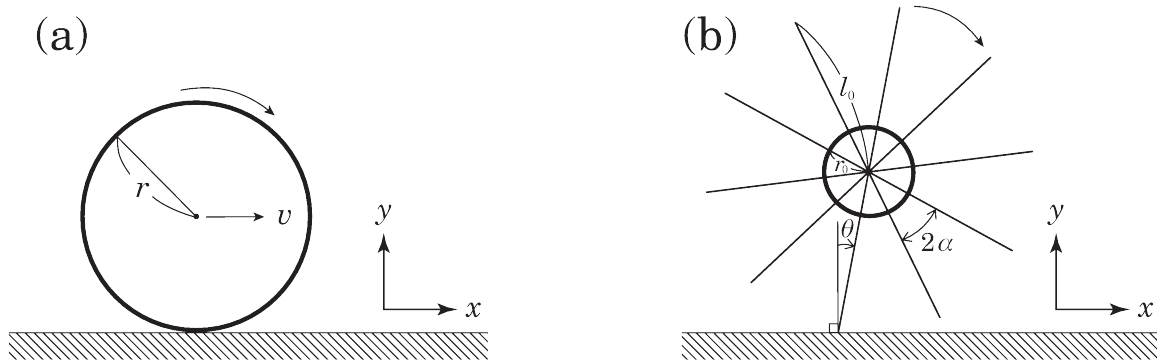


図 1

(問 1) 図 1(a) のような水平面上の円環の運動について，円環は平面上を滑らず，また摩擦によるエネルギーの損失なく転がる時，次の  から  の空欄にあてはまる値もしくは式を答えよ．

円環の中心まわりの慣性モーメント  $I_a$  を， $m$ ， $r$  を使って表すと，

$$I_a = \text{ア}$$

である．時刻  $t$  での円環の角速度を  $\omega(t)$  とする． $\omega(t)$  に関する運動方程式は，

$$I_a \frac{d\omega(t)}{dt} = \text{イ}$$

である．したがって， $t = 0$  における角速度を  $\omega_0$  とすれば，時刻  $t$  での角速度  $\omega(t)$  と円環の中心の  $x$  軸方向の速さ  $v(t)$  は，それぞれ，

$$\omega(t) = \text{ウ}, \quad v(t) = \text{エ}$$

である．

(問 2) 図 2(a) のような水平面上のリムレスホイールの運動において，支持脚は平面上を滑らず，また摩擦によるエネルギーの損失はないとする．また，遊脚が接地するとき，完全非弾性衝突が起こり，瞬時に遊脚から支持脚および支持脚から遊脚への遷移が起こるとする．支持脚が平面の法線方向と成す角を  $\theta$  (時計回りを正とする) とするとき，次の  から  の空欄にあてはまる値もしくは式を答えよ．

(次のページに続く)

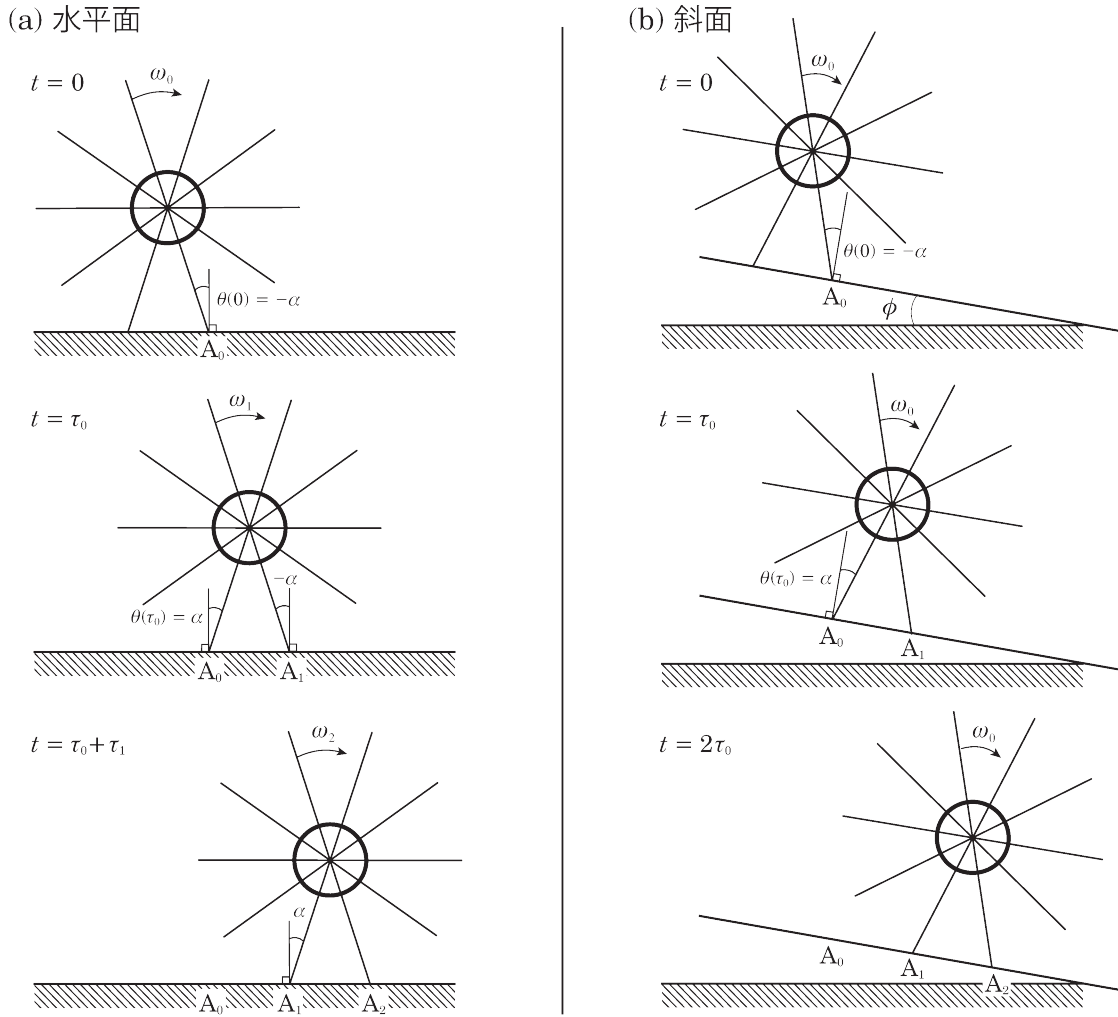


図 2

図 2(a) の上段のように、時刻  $t = 0$  において点  $A_0$  を支持脚の接点とする．点  $A_0$  まわりの慣性モーメント  $I_b$  を、 $m$ 、 $r_0$ 、 $l_0$  を使って表すと、

$$I_b = \boxed{\text{オ}}$$

となる．時刻  $t = 0$  において、点  $A_0$  のまわりに角速度  $\omega_0$  を与える．角速度  $\omega_0$  は、リムレスホイールが時計回りにしばらく回転するのに十分な大きさであるとし、遊脚から支持脚への遷移が起こる際の角速度の変化を考える．図 2(a) の中段のように、 $t = \tau_0$  で遊脚から支持脚への最初の遷移が点  $A_1$  で起こるとする．エネルギー保存の法則により、 $t = \tau_0$  の点  $A_0$  まわりの角速度は  $\boxed{\text{カ}}$  である．遊脚が点  $A_1$  に接地する直前において、点  $A_1$  をリムレスホイールの回転軸と考えたとき、点  $A_1$  まわりの重心の角運動量は  $\boxed{\text{キ}}$  であり、重心まわりの角運動量は  $\boxed{\text{ク}}$  である．このとき、リムレスホイールの全角運動量  $H_{(-)}$  は、

$$H_{(-)} = \boxed{\text{ケ}} \omega_0$$

である．一方、点  $A_1$  に接地した直後の点  $A_1$  まわりの角速度を  $\omega_1$  とすると、リムレスホイールの全角運動量  $H_{(+)}$  は、

$$H_{(+)} = \boxed{\text{コ}} \omega_1$$

(次のページに続く)

となる。したがって、角運動量保存の法則により、 $\omega_0$  と  $\omega_1$  の関係は、

$$\omega_1 = \boxed{\text{サ}} \omega_0$$

となる。さらに、リムレスホイールの回転に伴い遊脚が接地する点を順に、 $A_1, A_2, \dots, A_n, \dots$  とする。点  $A_n$  において遊脚が支持脚に遷移した直後の点  $A_n$  まわりの角速度を  $\omega_n$  とすると、

$$\omega_n = \boxed{\text{シ}} \omega_0$$

となる。

(問 3) 図 2(b) のように、リムレスホイールが水平面と角度  $\phi$  を成す無限に続く斜面を転がる運動を考える。前問で考えたように、リムレスホイールが水平面を転がる場合は、ステップ数  $n$  に対して角速度は指数関数的に減少する。しかし、斜面を下る場合は、重力の作用により 1 ステップあたりの平均速度が一定の運動が可能になる。このような運動について、次の  $\boxed{\text{ス}}$  から  $\boxed{\text{チ}}$  の空欄にあてはまる値もしくは式を答えよ。以下では、角度  $\alpha$  と  $\phi$  は十分小さく、 $\sin \alpha \simeq \alpha$ 、 $\cos \alpha \simeq 1$ 、 $\sin \phi \simeq \phi$ 、 $\cos \phi \simeq 1$  と近似できるものとする。

時刻  $t$  において点  $A_0$  を接点とする支持脚が斜面の法線方向と成す角を  $\theta(t)$  とするとき、 $\theta(t)$  についての運動方程式は、 $\theta$  と  $\phi$  について線形近似すると、

$$I_b \frac{d^2\theta(t)}{dt^2} = \boxed{\text{ス}}$$

となる。ここで、 $I_b$  は  $A_0$  まわりの慣性モーメントである。この運動方程式において、

$$\theta(0) = -\alpha, \quad \frac{d\theta(0)}{dt} = \omega_0$$

を満たす解は、

$$\theta(t) = \boxed{\text{セ}}$$

である。リムレスホイールの回転に伴い遊脚が接地する点を順に、 $A_1, A_2, \dots, A_n, \dots$  とする。点  $A_n$  から点  $A_{n+1}$  へ脚が運ばれる時間間隔が一定値  $\tau_0$  であるとき、遊脚が支持脚に遷移した直後の支持脚まわりの角速度が  $\omega_0$  であると仮定する。さらに、図 2(b) の上段のように、 $t = 0$  で支持脚が点  $A_0$  にあるとき、すでにこのような周期的な運動が実現されているとする。この運動状態が実現するためには、図 2(b) の中段のように遊脚が点  $A_1$  に接地する直前において、

$$\theta(\tau_0) = \alpha, \quad \frac{d\theta(\tau_0)}{dt} = \boxed{\text{ソ}} \omega_0$$

となる必要がある。これらの条件を満たす  $\omega_0$  は、

$$\omega_0 = \boxed{\text{タ}}$$

となる。したがって、リムレスホイールの平均の速さ  $\bar{v}$  は、

$$\bar{v} = \frac{2l_0 \sin \alpha}{\tau_0} \simeq \frac{2l_0 \alpha}{\tau_0} = \boxed{\text{チ}}$$

となる。

## [Ⅱ - 3]

画像を用いた距離計測に関連する以下の問いに答えよ。

(問 1) カメラ 2 台を光軸が平行になるように並べ、二眼ステレオ法により物体までの距離計測を行うパッシブステレオ法について考える。

2 台のカメラの光軸は  $z$  軸と平行とする。カメラに入射する全ての光線は、それぞれのカメラの光学中心  $O_L(-b/2, 0, 0)$  および  $O_R(b/2, 0, 0)$  を通る (ピンホールカメラモデル)。撮像面は、一般的には  $z < 0$  の領域に配置されるが、像が反転してわかりにくいいため、撮像面を各々の光学中心に対して点対称に折り返し、図 1 のように平面  $z = f$  の上に配置する。撮像面  $S_L$  および  $S_R$  の中心は、それぞれ  $(-b/2, 0, f)$  および  $(b/2, 0, f)$  である。  $b$  は基線長、  $f$  は焦点距離と呼ばれ、いずれも正の値である。光学中心  $O_L$  と撮像面  $S_L$  の組、および、光学中心  $O_R$  と撮像面  $S_R$  の組から構成されるカメラを、それぞれカメラ  $C_L$  およびカメラ  $C_R$  と呼ぶ。

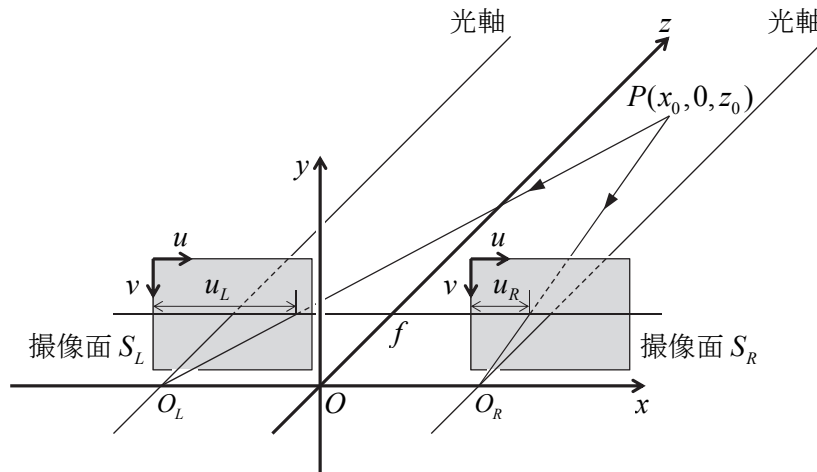


図 1

- (ア) ピンホールカメラモデルでは、空間内の点は、点とカメラの光学中心とを結ぶ直線と、撮像面との交点に投影される。撮像面  $S_L$  および  $S_R$  の大きさを幅  $d_w$ 、高さ  $d_h$  とし、それぞれの撮像面上に、図 1 のように二次元座標系  $uv$  を取る。点  $P(x_0, 0, z_0)$  が、撮像面  $S_L$  および  $S_R$  に投影される点の  $u$  座標  $u_L$  および  $u_R$  を、  $x_0, z_0, b, f, d_w$  を用いて表せ。
- (イ) ある点が二つの撮像面それぞれに投影された際の、撮像面上での位置の差を視差  $D$  とし、  $D = u_L - u_R$  と定義する。対象物までの距離  $z_0$  を視差  $D$  から求める式を書け。
- (ウ) 一般的に、画像から求めた視差には、計測に伴う誤差が含まれる。視差に誤差  $\Delta D$  が生じた際に、距離の真値が  $z = z_0$  である点の  $z$  座標 (奥行き) の推定値に生じる誤差を、  $b, f, z_0, \Delta D$  を用いて表せ。ただし、  $D \gg \Delta D$  とする。
- (エ) パッシブステレオ法の奥行き分解能を向上させるために考えられる方法と、その方法により分解能が向上する理由について述べよ。また、そのときに生じる欠点について述べよ。ただし、撮像面に配置される撮像素子の性質 (解像度および画素数) は変更できないものとする。

(次のページに続く)

(オ) ヒトが網目のような規則的な模様を持つ物体を見た場合に、実際の距離とは異なる距離に、物体があるように感じられることがある。この理由を、パッシブステレオ法の原理に基づき、図を用いて説明せよ。

(問 2) 問 1 (オ) で見られるパッシブステレオ法の問題点を解決する方法として、以下に述べるアクティブステレオ法がある。

まず、問 1 のカメラ  $C_L$  を取り除き、点  $O_L(-b/2, 0, 0)$  に点光源を、撮像面  $S_L$  があった位置にスライド面  $S'_L$  を置く。スライド面  $S'_L$  の中心は  $(-b/2, 0, f)$  である (図 2)。スライド面上には投影パターンが設定され、光が透過できる領域を限定する。点光源から発してスライド面を通過した光は、スライド面上に設定された投影パターンを、計測対象に投影する。点光源とスライド面の組によって構成される投影装置をプロジェクタと呼ぶ。プロジェクタによりパターンが投影された計測対象を、カメラ  $C_R$  で撮像し、スライド面  $S'_L$  上の位置とカメラ  $C_R$  の撮像面  $S_R$  上の位置を対応付けることにより、パッシブステレオ法と同じ原理に基づく距離計測が可能となる。

(ア) スライド面の大きさを幅  $d_w$ 、高さ  $d_h$  とし、スライド面上に、図 2 のように二次元座標系  $uv$  を取る。

スライド面は格子状の画素に区切られており、画素単位で投影パターンを設定できる。横方向の画素数を  $W$ 、縦方向の画素数を  $H$  とする。スライド面左上の画素を画素  $(0, 0)$  とし、その画素から数えて横方向に  $i$  画素、縦方向に  $j$  画素進んだ位置にある画素を画素  $(i, j)$  と表記する。図 2 左上にスライド面  $S'_L$  上の画素配置を示す。このとき、画素  $(i, j)$  の中心の  $uv$  座標を求めよ。

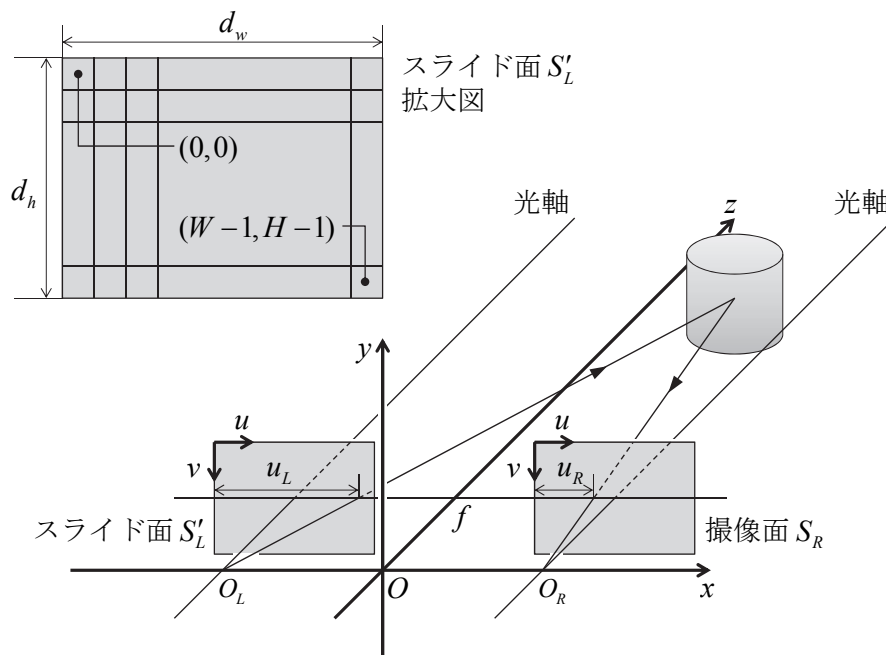


図 2

(次のページに続く)



(イ) スライド面の横方向の画素数  $W$  および縦方向の画素数  $H$  を 256 とする. 計測対象に対して, 画素 (163,128) のみを光が通過するスポット光パターンを投影したところ, カメラ  $C_R$  の光軸上の点が明るくなった (すなわち  $u_R = d_w/2$  である). 撮像面およびスライド面の幅  $d_w$  が 15mm, 焦点距離  $f$  が 20mm, 基線長  $b$  が 100mm のとき, スポット光が投影された点の  $z$  座標を求めよ.

(ウ) スライド面の縦一列の画素のみを透過して投影されるパターン光を, スリット光と呼ぶ. 図 3 上に示す 5 種類のスリット光を計測対象に投影し, カメラ  $C_R$  で取得した画像を図 3 下に示す. これらの結果から, 計測対象物体の形状についてわかることを述べよ.

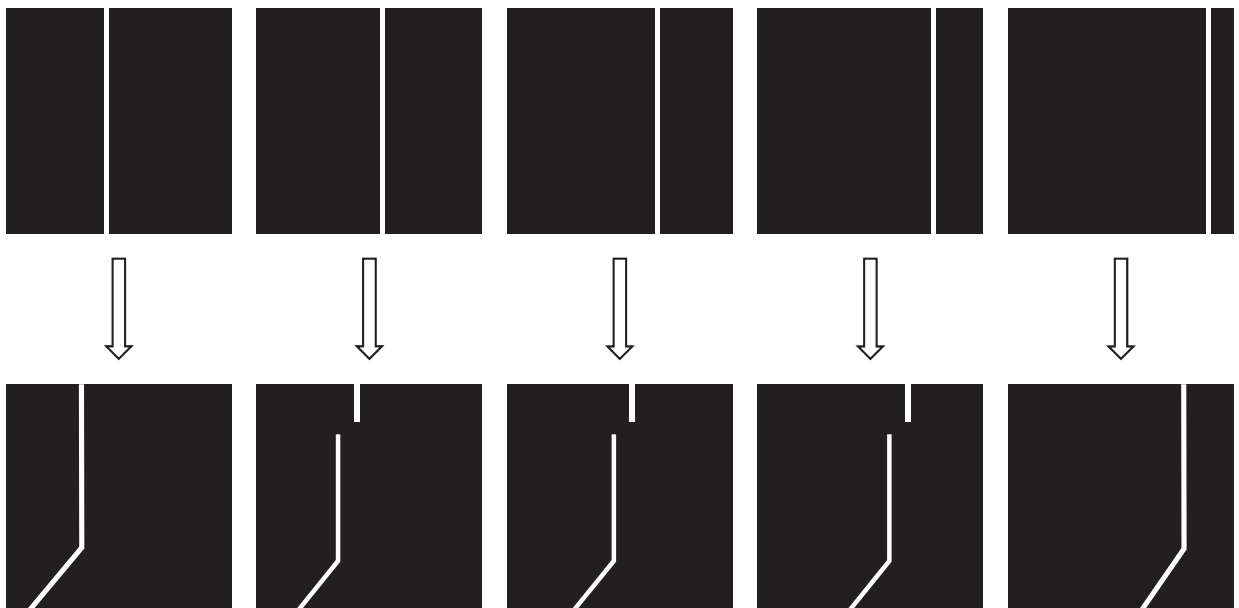


図 3

(問 3) ヒトが視覚から周囲の空間の奥行きや三次元構造を把握する上では, 様々な手掛かりを用いている. 前問までで取り扱ったステレオ法は両眼視差の利用に相当する. 奥行き感をもたらす視覚系の手掛かりを両眼視差以外に三種類挙げ, それぞれが奥行きの把握にどのように寄与しているかを説明せよ.